Statistical signal processing for echo signals from ultrasound linear and nonlinear scatterers

Yan Yan

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
December 22, 2009

Abstract

In diagnostic ultrasound contrast imaging (UCI), changes in the microvasculature due to some high mortality diseases can be detected so as to provide an early screening tool for treatment monitoring. Ultrasound contrast agents (UCAs), which are introduced in the field of UCI, have attracted large amounts of attention and research since they were discovered. Most of existing investigations about UCAs focus on optical observation and theoretical prediction, which are inadequate and questionable. In terms of experiments-based research on UCAs, although more accurate experimentally measured UCA responses can be obtained, there is a lack of advanced signal processing tools for the analysis of the echo signals. Therefore, the development of new signal processing techniques is required in future strategies in medical care. This thesis develops statistical signal processing tools for the analysis of characteristics of ultrasound echo signals from UCAs and soft tissue based on *in-vitro* experiments. These characteristics will result in discrimination of responses from UCAs and tissue.

The contributions of this thesis can be generalized as proposing a *parametric* signal processing method for the measured ultrasound echo signal analysis within a Bayesian framework. Although various general signal processing techniques have been demonstrated to be very useful in many research fields, the results of their applications to ultrasound field are not satisfactory. This is mainly because the signals under investigations are all based on experimental measurements and there is no such signal processing tools that are particularly designed for the analysis of ultrasound echoes. In this thesis, in order to extract the features of ultrasound echo signals, a *parametric* model is proposed for the measured echo signals and its parameters are estimated in a statistical way. This signal model contains information on both the time domain and the frequency domain, and also takes account of the transducer that has a band-limited property.

Furthermore, a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm is employed and modified to obtain samples from non-linear posterior distributions of the desired model parameters, which are able to provide corresponding estimates for the parameters. The modification of the algorithm is made in particular to incorporate the transducer characteristics and to extract characteristics of the echo signals in the time and the frequency domains simultaneously. The newly proposed signal model and the modified estimation algorithm are able to generate better results than conventional signal processing methods. Also, from the ultrasound imaging point of view, more knowledge about the behavior of ultrasound scatterers can be learnt and thus a direction of discriminating the ultrasound echoes can be pointed out.

Acknowledgements

Finally I come to the end of this long journey of my four-year PhD life at the University of Edinburgh. Now, I would like to take this opportunity to express my appreciation to a number of people who have provided me with invaluable help over the past years.

First of all, I would like to give my sincere thanks to my supervisor Dr. James Hopgood for his kindness help and guidance during my research and for proofreading this thesis. Without his continuous help and encouragement, it would not possible for me to complete the work.

Moreover, I am also grateful to Dr. Vassilis Sboros, who introduced a totally new area of ultrasound contrast imaging to me and provided me enumerate results of experimental measurements. It is a very important part of my research and a great part of this thesis has also been peer-reviewed by him.

Furthermore, I would like to thank Prof. Steve McLaughlin for offering me a studentship to support my research. The financial aid I received from Biologically inspired acoustic system (BIAS) project is also appreciated and acknowledged.

During the past four years in Edinburgh, I have made lots of friends and had a wonderful time in my life abroad. Many people deserve warm thanks, especially Hao Wang, who gave me strong support during the period of thesis writing-up.

Lastly and the most importantly, I wish to give my special thanks to my parents. No words can describe my deepest affection and gratitude to them. Their endless and priceless love and encouragement make me hold on to the end of this tough journey.

This thesis is dedicated to my parents and I hope I have not disappointed them.

Yan Yan

Edinburgh, December 2009.

Contents

List of figures

List of tables

Acronyms and abbreviations

AIC	Akaike information criterion
AM	amplitude modulated
ANN	artificial neural network
AR	autoregressive
ARMA	autoregressive moving average
BIC	Bayesian information criterion
CAT	criterion for autoregressive transfer function
CT	computed tomography
CTR	contrast to tissue ratio
CWT	continuous wavelet transform
DAG	directed acyclic graph
dB	decibel
DFT	discrete Fourier transform
DWT	discrete wavelet transform
EMD	empirical mode decomposition
ESM	exponential sinusoidal model
ESPRIT	estimation of signal parameters via rotational invariance techniques
FPE	final prediction error
FT	Fourier transform
GASM	Gaussian-apodized sinusoidal model
GIC	generalized information criterion

HF	high frequency
HQ	Hannan-Quinn
HT	Hilbert transform
HTWD	Hilbert transform-wavelet denoising
IDFT	inverse DFT
i.i.d.	identically independent distributed
IMF	intrinsic mode function
KL	Kullback-Leibler
LF	low frequency
LMS	least mean squares
LS	least squares
MAP	maximum a posteriori
MB	microbubble
MCMC	Markov chain Monte Carlo
MDL	minimum description length
MH	Metropolis-Hasting
MI	mechanical index
ML	maximum likelihood
MMSE	minimum mean squared error
MRA	Mallat's multiresolution analysis
MRI	magnetic resonance imaging
MSE	mean squared error
MUSIC	multiple signal classification

PDF	probability density function
PET	position emission tomography
PHD	Pisarenko harmonic decomposition
PI	pulse inversion
PM	power modulation
PIAM	pulse inverted amplitude modulation
PSD	power spectral density
PVDF	Polyvinylidenfluorid
PZ	pole-zero
rjMCMC	reversible jump MCMC
RLS	recursive least squares
ROI	region of interest
RTF	receiver transfer function
SD	standard deviation
SCS	solid copper sphere
SNR	signal-to-noise ratio
STFT	short-time Fourier transform
SURE	Stein's unbiased risk estimator
TMM	tissue-mimicing material
UCA	ultrasound contrast agent
UCI	ultrasound contrast imaging
VAD	voice activity detection
WGN	white Gaussian noise
WT	wavelet transform

Nomenclature

$\otimes$	convolution operation
$< \cdot >$	inner product
$\| \cdot \|$	Euclidean norm
$\mathbf{I} \cdot \mathbf{I}$	absolute value
$\mathbf{E}[\cdot]$	expectation operation
$x(t)$	AM signal
$\hat{x}(t)$	HT of $x(t)$
$x_a(t)$	analytic signal of $x(t)$
$s(t)4$	real baseband signal
P_{s_j}	power of the signal $s(t)$ with respect to the j^{th} component
P_{n_j}	power of the noise $n(t)$
$\phi(t)$	mother wavelet
$\phi_{a,b}(t)$	child wavelet
WT_ϕ	wavelet coefficients
η_T	threshold function
$S_{m,n}$	wavelet coefficients used in MRA
$\text{sgn}(\cdot)$	sign function
f_l	spectral bin
$\phi_k(f_l)$	SNR measure used in VAD technique
$P_{xx,k}(f_l)$	signal PSD of the current frame k
$\hat{P}_{vv}(f_l)$	estimated noise PSD
$\eta_{T,k}(f_l)$	adaptive threshold used in VAD technique
$p(\cdot)$	PDF
$Pr(\cdot)$	probability function
$D(k)$	decision vector used in VAD technique
$w(n)$	window function
$x_w(n)$	windowed signal
$\hat{X}_N(k)$	DFT of the windowed signal $x_w(n)$
$\hat{R}_x(e^{j\omega})$	periodogram of the signal $x(n)$

$\hat{r}_x(l)$	autocorrelation sequence of the signal $x(n)$
$\hat{V}(k)$	N-point DFT of the windowed signal
$\hat{R}_x(k)$	periodogram at discrete frequencies
$w_k(n)$	orthogonal data tapers
$\hat{R}_x^{(MT)}(e^{j\omega})$	estimated multitaper spectrum
$X(m,\omega)$	STFT of the signal $x(n)$
k	model order
$\hat{\sigma}_k^2$	modeling error with model order of k
$\mathbf{y}$	observation vector with noise
$I_{N\times N}$	$N \times N$ identity matrix
α	positive constant used in GIC
C	constant used in HQ
$\mathcal{M}_k$	a set of candidate models
Ψ	a vector of signal model parameters
$K_B(\cdot)$	Bayes factor
N	number of observations
$\mathbf{s}$	noise-free signal
$\mathbf{A}_k$	amplitude vector
$\mathbf{R}_k$	k^{th} basis function
$\sin(\cdot)$	sine function
$\cos(\cdot)$	cosine function
$\exp(\cdot)$	exponential function
$a_{c,k}$	amplitude coefficient for the cosine function at the k^{th} component
$a_{s,k}$	amplitude coefficient for the sine function at the k^{th} component
E_k	energy of the k^{th} component in the signal
$d(t)$	decaying function
$w(t,\alpha)$	Gaussian window with a standard deviation of α
T_i	the i^{th} change-point
Y_0	noise-only event
Y_1	a single pulse segment event
Y_{k_m}	m pulse segments event
$\boldsymbol{\omega}_k$	frequency values for the k^{th} component
$S_i(t)$	noise-free signal for the i^{th} pulse segment

$\triangle f$	frequency resolution
T_s	time duration
$Var(f)$	variance of the peak in frequency spectrum
$\mathbf{D}$	the matrix in the full length signal
$\mathbf{a}_{k_m}$	linear coefficients for the matrix $\mathbf{D}$
$\boldsymbol{\omega}_{k_m}$	frequency values in all m pulse segments
$\mathbf{T}_{2m}$	change-points in the full length signal with m pulse segments
σ_n^2	noise variance in the full length signal
$\boldsymbol{D}_i$	the matrix containing frequency information in the i^{th} pulse segment
$\boldsymbol{a}_{k_i}$	amplitudes in the i^{th} pulse segment
$\sigma_{k_i}^2$	noise variance in the i^{th} pulse segment
N_i	number of observations in the i^{th} pulse segment
k_i	number of frequency in the i^{th} pulse segment
$\mathbf{y}_i$	the observations in the i^{th} pulse segment
$\mathbb{I}_\Omega(\omega)$	if $\omega \in \Omega$, it equals 1, otherwise it equals 0
N_f	number of frequency bins
$\sigma_{RW_f}^2$	variance of the random walk for a frequency value
b_k	probability of choosing a *birth* move
d_k	probability of choosing a *death* move
u_k	probability of choosing an *update* move
$\boldsymbol{\psi}_k$	all model parameters in the full length signal with model order k
$\boldsymbol{\theta}_{k_i}$	all model parameters in the i^{th} pulse segment with model order k
$\mathcal{N}(\cdot)$	normal distribution
$\mathcal{G}a(\cdot)$	Gamma distribution
$\mathcal{IG}(\cdot)$	inverse Gamma distribution
$\Gamma(\cdot)$	gamma function
σ_T^2	variance of updating a single change-point
$\mathbf{s}_r^{(iter)}$	reconstructed noise-free signal at the current iteration *iter*
$\mathbf{e}$	residual error between the original signal and the reconstructed signal
$P_{exp}(t)$	pressure of the measured ultrasound scattering in the time domain
$P_{theory}(t)$	pressure of the theoretical ultrasound scattering in the time domain
$\boldsymbol{\beta}$	polynomial coefficients
P_{NG}	noise gain

$H(z)$	a system denoted in Z-plane
$h(n)$	a system in the time domain
$F_{input}(\omega)$	the FT of the input signal to the ultrasound receiver
$F_{output}(\omega)$	the FT of the output signal to the ultrasound receiver
$\mathbf{H}$	receiver matrix
$\mathbf{h}$	impulse response of the ultrasound receiver
$\hat{\mathbf{H}}$	frequency response of the ultrasound receiver
σ_m^2	uncertainties of polynomial coefficients of magnitude response
σ_p^2	uncertainties of polynomial coefficients of phase response
M	length of the ultrasound receiver in the time domain
L	length of the filtered signal observations

Chapter 1

Introduction

Nowadays, since the introduction of ultrasound contrast agents (UCAs) to diagnostic ultrasound contrast imaging (UCI) techniques, many investigations have focused on theoretical prediction, which are inadequate and questionable. In order to characterise the measured ultrasound echo signals based on *in-vitro* experiments, general signal processing approaches cannot meet the requirements. Therefore, more advanced signal processing tools especially designed for the analysis of echo signals need to be developed. This thesis is dedicated to extracting characteristics of ultrasound echo signals from tissue-mimicking materials and UCAs using statistical signal processing algorithms. This introductory chapter first gives a brief review of the development of ultrasound contrast imaging techniques in Section 1.1. Then the origination and motivation of the study is presented in Section 1.2. Moreover, the objective and major contributions are stated in Section 1.3. Finally, Section 1.4 provides an overview of the structure of the whole thesis.

1.1 Development of ultrasound contrast imaging techniques

Ultrasound diagnostic techniques have been shown to be powerful, versatile and well suited to medical practice [4]. Ultrasound imaging is an important means of obtaining information about the organ structures, especially the cardiovascular function of the body with the employment of ultrasound. Nowadays, the new idea of intravascular injection of gas-containing particles for information gathering becomes more and more popular in advanced ultrasound imaging [5]. The introduced particles in tissue, termed as UCAs cause changes that can be derived from the ultrasound echo returns, such as total intensity, spectral contents and some temporal changes. By detecting the presence of UCAs and differentiating them from tissue, the resolution and sensitivity of ultrasound images can be greatly improved.

1.1.1 Diagnostic ultrasound

Ultrasound is the transmission of mechanical vibrations through matter and is usually termed as the sound which has very high pitch that cannot be heard by humans [4]. The frequency of ultrasound is represented by the number of particle oscillations per second when ultrasound is propagating, and is measured by converting mechanical oscillations into voltage oscillations. Ultrasound imaging is one of the most important applications of ultrasound, which can produce images of soft tissue structure. The ultrasound is transmitted into the body and the echoes are detected by reflections at tissue boundaries. In this way, ultrasound images are created. According to these images, the size and the nature of tissue structures can be determined. During the past several decades, ultrasound imaging is more and more popular in diagnosing all vascular related diseases, including cancer and heart diseases.

Since 1990's, UCAs have been employed in ultrasound imaging [6]. They are injected intravenously in an attempt to detect and quantify the tissue perfusion by examining the ultrasound echoes from tissue and UCAs. They provide improvements to other advanced imaging techniques, such as position emission tomography (PET) [7], magnetic resonance imaging (MRI) [8] and computed tomography (CT) [9] imaging of perfusion. As these techniques have disadvantages of high toxicity levels, leakage from the vascular system and a high cost of repetitive imaging, ultrasound imaging provides a non-destructive, non-invasive and cost-efficient paradigm for medical diagnosis [10].

1.1.2 Ultrasound contrast microbubbles

Early in the late 1960's, UCAs were first discovered accidentally by cardiologist Charles Joiner, who injected indocyanine green dye into a heart. The observed increasing echogenicity of the blood containing the dye and the small air bubbles formed on the catheter tip produced transient echoes from the ultrasound [11]. Until 1990's, the use of microbubbles (MBs) as UCAs had been widely used in bio-medical research [6].

Current UCAs are composed of gas-filled capsulated microspheres small enough to go through micro-circulations in human bodies. The average diameter of the ultrasound contrast MBs ranges between 2μm and 7μm. Encapsulation is necessary to prevent rapid dissolution of the gas content into the blood [12]. When exposed to ultrasound, a contrast MB starts to oscillate under the pressure of the sound field. This oscillating behaviour results in high scattering strength of the contrast MB [13]. Com-

pared to soft tissue, MBs are more compressible and expandable when insonificated with ultrasound. Therefore, in the ultrasound literature, soft tissue are termed as linear scatterers, which have spectra with dominance at the fundamental frequency of the incident wave. By contrast, MBs are usually termed as nonlinear scatterers in terms of incident frequency as the spectral contents in their echo signals have more harmonics than those in the transmit pulse wave [5, 6, 10].

A significant amount of research has focused on the development of a theoretical model, which describes the oscillation of a MB as a function of the incident sound wave [14–17]. The major difference between these models lies in the modeling of shells of MBs according to different properties of shell materials. However, the exact properties of MBs and layout of their shells are not well understood. Since 2003, with the help of very high frame rate cameras, it is possible to observe the oscillations of contrast MBs optically [18] and to study single MBs rather than clouds of MBs experimentally. This provides a basis for better understanding of the MB behaviour and for designing more accurate models for MBs.

1.1.3 Excitation schemes for the detection of MBs

The usability of contrast MBs in UCI depends mostly on the ability to detect their presence in blood and being able to distinguish them from tissue [19]. In theory, the echo returns from tissue have one dominant fundamental frequency as the spectrum of incident pulse exhibits. However, the echo returns from MBs have more harmonics due to their oscillations (including contraction and compression) exposed to ultrasound. Therefore, most signal processing techniques in the ultrasound literature concentrate on exploiting the strong nonlinear characteristics of MBs. The excitation schemes for transmit pulse examined nowadays are designed to maximise the difference in their harmonic frequency contents and thus increase the contrast to tissue ratio (CTR) and the signal-to-noise ratio (SNR) of the ultrasound echo signals. They can be divided into three categories:

- Passive technique: based on a single excitation pulse, e.g. harmonic imaging.
- Active multi-pulse technique: based on multiple pulse sequences to improve the CTR, e.g. pulse inversion, power modulation, chirp excitation, etc.
- Destruction-based detection of MBs due to high mechanical index (MI).

These schemes will be investigated in Section 2.1 in more detail. Most of the techniques use a spectral filtering approach to extract the harmonics and separate the respective contribution from soft tissue and contrast MBs.

1.2 Motivations

The ultimate objective of analysing MB behaviour in UCI is to differentiate MB responses from tissue responses based on experimental measurements rather than theoretical assumptions. Nevertheless, either from the perspective of ultrasound imaging or from the signal processing point of view, some problems will arise, which will be discussed as follows.

1.2.1 Limitations of theoretical models

In the ultrasound imaging literature, most of the techniques for improving UCI focus on designing new excitation pulses in theory. By transmitting an incident waveform using the developed pulsing schemes, the CTRs of echo signals at fundamental and harmonics can be theoretically calculated. Due to the nonlinear signature of MBs, the CTRs of the theoretical echo signals at even harmonics, especially the second harmonic, can be improved to a large extent.

Nevertheless, there is limited knowledge on exactly how MBs behave within the tissue and how the reflected ultrasound signals should be interpreted. Although there are a variety of investigations about MB modeling in [20–24], they are not adequate for explaining actual MB behaviour due to two main reasons. First, the validity of models has not been verified experimentally; Second, there are some certain assumptions that the gas is perfect and no instability of fast moving liquid-gas interface exists for the ease of mathematical calculation. As a consequence, the only way to correctly understand MB behaviour is through experimental measurements, thus analysing the real behaviour of MBs based on actual measurements is required. All work in this thesis is based on measured ultrasound echo signals instead of theoretical assumptions.

1.2.2 Advantages of single MBs used in experimental measurements

In many existing experimental measurements, the analysis of MBs are usually carried out in suspension and are studied in dense populations. However, if the concentration of MBs is relatively high, although it is not observable, it might affect measurements in

a complex manner. Sboros [25] pointed out limitations of the measurements from MB populations. The most common problems are multiple scattering effects and bubble interactions. The effect of multiple scattering is not dominant but the assumption that it can be negligible has not been justified. Moreover, in a high density suspension, MBs are close enough to affect oscillatory motion of each other and thereby display different acoustical behaviour, compared to the behaviour of isolated MBs. Accordingly, the experimental measurements, as well as the theoretical predictions, of scattering from single MBs need to be examined.

The measured ultrasound echo signals under investigation in this thesis are all acquired from single scatterers, especially single MBs. Therefore, the analysis results are more convincing compared to those from dense populations.

1.2.3 Deficiency of conventional signal processing techniques used in ultrasonics

Existing development of UCI techniques concentrate on improving the excitation pulsing schemes for transmit pulses. However, they have been intensively studied during the past decades. Moveover, most of the techniques are only justified using theoretical prediction, instead of using experimental measurements.

Leaving aside the evolution of designing excitation schemes, the underlying characteristics of ultrasound echo signals from various scatterers is another perspective that needs explored. In order to investigate these characteristics, more advanced signal processing techniques need to be employed and developed for studying experimentally measured echo signals, which are helpful in learning the true behaviour of scatterers, especially MBs.

1.2.3.1 Limitation of matched filters

In most of the popular transmit pulsing design schemes, including the passive single pulse excitation and the active multi-pulse schemes, an optimal filter is required to extract the fundamental frequency component and the harmonics respectively, and thus to calculate their CTRs. If the transmit waveform propagates through a linear medium, and the only interference is white Gaussian noise (WGN), the optimal receiver is a matched filter. Matched filters were first derived by North [26] in 1943 and now are often used in radar and sonar to detect deterministic signals in the presence of additive white Gaussain noise [27].

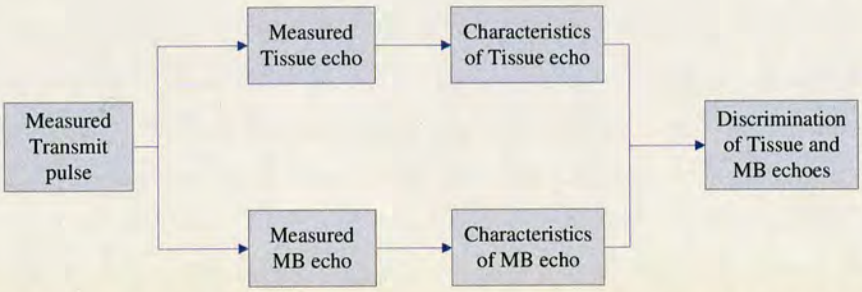


Figure 1.1: *Discrimination procedure based on the measured echo returns.*

While in the application of ultrasound contrast imaging, the medium, such as MBs, is nonlinear, the clutter is non-Gaussian, and the medium characteristics are unknown. The matched filter cannot be used for receiver optimisation any more. Therefore, some advances in signal processing techniques should be incorporated into characterising measured echo signals from both tissue and MBs.

1.2.3.2 A joint estimation framework

As the time domain information and the frequency domain information are both of significant importance in differentiating various responses, an effective way to characterise ultrasound echo signals should be performed in a single unified framework. Nevertheless, traditional signal processing techniques based on *non-parametric* methods in analysing these responses are usually carried out in separate domains and individual steps. It is sub-optimal to have a two-step estimation scheme. Furthermore, the CTRs are obtained individually and thus cannot be utilised to discriminate the distinct responses. If a unified estimation framework is available, the joint estimation method will make it much easier to extend the analysis of the echo signals to the automatic discrimination of them. The procedure of the differentiation can be illustrated in a diagram in Figure 1.1.

Therefore, in order to provide a way of estimating characteristics of measured ultrasound scatterers in a single joint framework, *parametric* methods will be introduced. An appropriate signal model is first proposed and parameter estimation algorithms are then developed. In this way, ultrasound echo signals can be characterised by estimating the signal model parameters in both the time and the frequency domains.

1.2.3.3 Reveal real MB behaviour

As mentioned in Section 1.2.1 and Section 1.2.2, the echo signals under analysis in this thesis are based on experimental measurements. In actual measurements, a membrane hydrophone was used to calibrate the transmitted acoustic field [3]. However, since the exact centering of MBs in the transmitted field is very difficult and cannot be done with a precision less than 1mm from the central line, it is much more complex to do the calibration for a received beam from single MBs using the same hydrophone. Therefore, an ultrasound transducer, which was used for transmitting a pulse waveform, was also used as a receiver for the echo signals from ultrasound scatterers.

Nevertheless, the transducer has a band limited property. It will have an influence on the estimation of spectral contents of scatterers, which may lead to inaccurate understanding of their behaviour, especially for MBs. Previous investigations in the ultrasound literature usually used the transducer spectral sensitivity to normalize the received signals [28] or use a so-called perfect reflector as a means of normalizing backscattered signals [15,29,30]. However, in the first method, the transducer spectral information cannot be easily obtained; the second method is not suited to the nonlinear scatterers, e.g. MBs.

In order to take the transducer characteristics into consideration, the diagram in Figure 1.2 illustrates an intuitive way to recover the true spectra of MBs and thus reveal real MB behaviour from the signal processing point of view. Specifically, the characteristics of the transducer, especially its spectrum, are first examined by comparing the theoretical and the measured echo signals from linear scatterers. Then by combining the received experimental measurements of MB echoes with the calibrated transducer characteristics, real MB behaviour can be revealed. In this process, how to combine the transducer characteristics and the measured MB responses is a critical issue. This requires many advanced signal processing techniques, which can provide powerful algorithms for inference, estimation and detection.

1.3 Objective and contributions

1.3.1 Objective

This thesis aims to jointly extract the characteristics of experimentally measured echo signals from various ultrasound scatterers in the time domain and the frequency domain. Since most of the research focused on MB behaviour according to theoretical

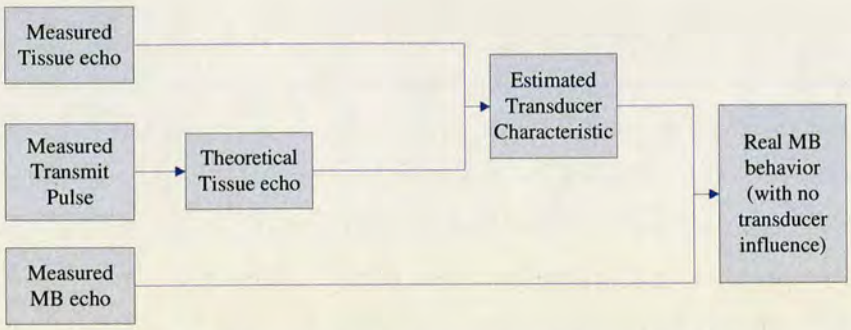


Figure 1.2: *The procedure of extraction of the real MB behaviour.*

models, which usually do not completely describe the MB movement, this study aims to discover the characteristics of MB echo returns entirely based on actual measurements. For ease of problem solving, the proposed estimation approaches are tuned to one particular measurement, which will be discussed in Chapter 2. In order to meet the goals of studying ultrasound echo signals from single MBs, and differentiating their responses from tissue, an estimation system is illustrated in Figure 1.3. It can be divided into two parts. One is estimation based on *non-parametric* approaches, which will be examined in Chapter 3. The other part is based on *parametric* approaches. It proposes a signal model for ultrasound echo signals and then develops several algorithms to estimate the model parameters. Details can be found in Chapter 4 and Chapter 6. Furthermore, the transducer characteristics are also incorporated into the proposed model, which will be addressed in Chapter 7. In this way, characteristics of the ultrasound echo signals can be extracted, which leads a possible way to discriminating the echo returns from MBs and soft tissue.

1.3.2 Key contributions

The major contributions in this thesis can be summarized from two perspectives. One is from ultrasound contrast imaging point of view and the other is from signal processing point of view.

1.3.2.1 Contributions to ultrasonics

Due to insufficient knowledge about signal processing techniques, researchers specialised in ultrasonics usually use the most traditional approaches and the very basic signal processing techniques to deal with the obtained signals. In this thesis, some effective signal processing approaches, which have seldom been applied in the ul-

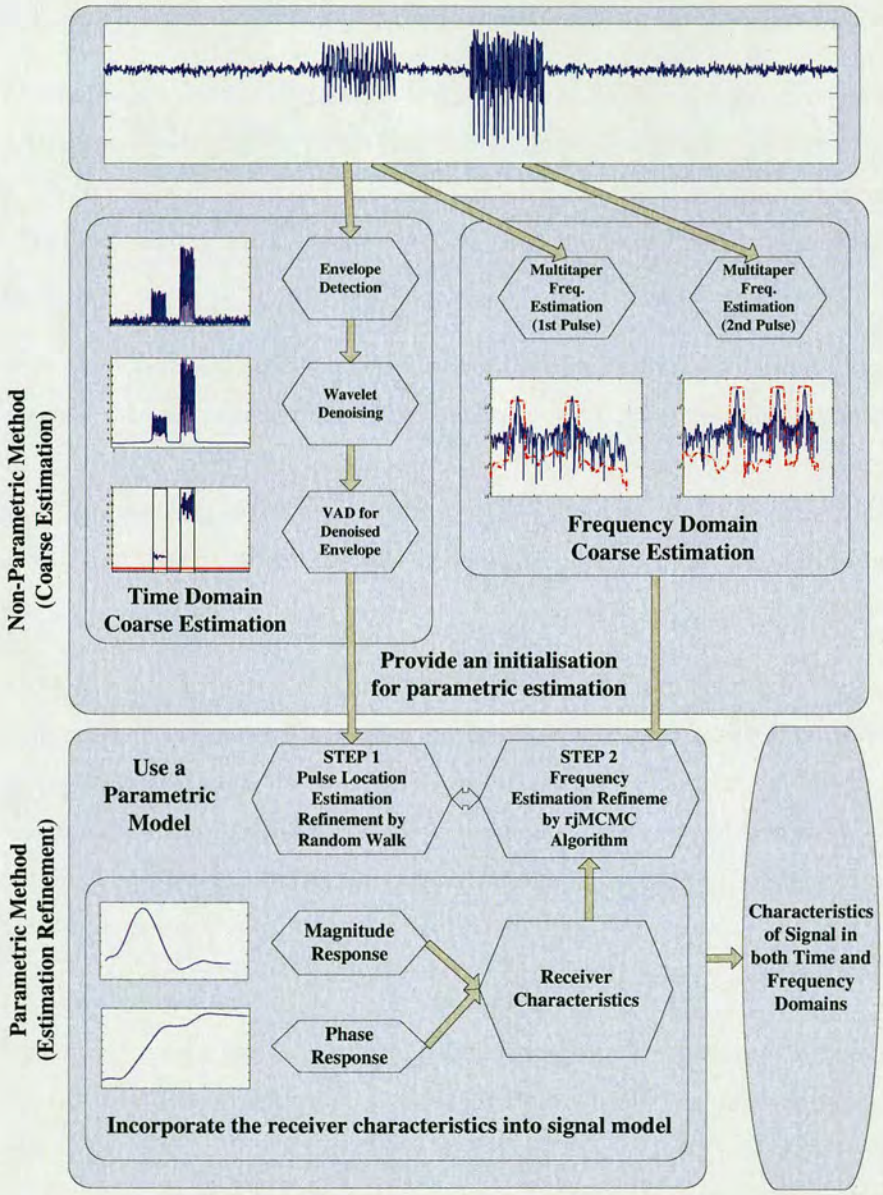


Figure 1.3: *The diagram for the complete estimation for ultrasound echo signals.*

trasound literature yet have been widely used in other applications, are introduced. Moreover, some new estimation and detection algorithms are developed, particularly for the analysis of ultrasound echo signals.

- A novel *non-parametric* method for estimating pulse locations in signals is proposed in the time domain, especially for analysing a signal with multiple pulses, e.g. MB echoes. It combines the envelope detection method using Hilbert transform and voice activity detection technique, which takes advantages of two ap-

proaches and is superior to each single method.

- A *parametric* inference method for spectral estimation within a Bayesian framework is employed in the frequency domain for analysing the measured ultrasound echo signals. It outperforms the traditional Fourier analysis based methods with more accurate estimation as some inherent limitations of them have been pointed out in [31,32].

- A signal model specially designed for the measured ultrasound echo signals is proposed, which contains both temporal and spectral information. Then a new sampling strategy is developed for the sophisticated posterior distribution of model parameters. In this way, an estimation system is created and the measured ultrasound echo signals can be characterised automatically in a single joint framework.

- Taking account of transducer influence on the measured ultrasound echo signals, an improved signal model which incorporates the transducer characteristics as part of the model parameters is proposed. Moreover, a sampling strategy specially designed for the corresponding posterior distribution is also developed. Therefore, the true spectra of the measured ultrasound responses from various scatterers are recovered, especially for MBs. This will help to learn more about the MB behaviour in UCI.

- The estimated temporal and spectral information of the measured ultrasound echo signals from soft tissue and MBs point a possible direction for discrimination between their responses. Moreover, it will assist in designing new adaptive excitation schemes for the incident pulse waveforms in future work.

1.3.2.2 Contributions to signal processing algorithms

From the signal processing point of view, the contributions can be summarized as follows:

- Most of the time-frequency analysis algorithms in the literature are based on *non-parametric* methods, e.g. short-time Fourier transform (STFT) and wavelet transform. While most of *parametric* spectral estimation methods do not localize in time, such as [33, 34]. This thesis suggests a new way to look at the time-

frequency analysis using a *parametric* method within a Bayesian framework. It can extract the temporal and spectral contents of a signal simultaneously.

- A new signal model for the measured ultrasound echo signals is proposed, which includes both the time domain information and the frequency domain information. Moreover, a new strategy for sampling from the derived posterior distribution of the model parameters, is developed. More accurate results can be generated than conventional signal processing techniques that are commonly used in ultrasonics.

- By taking transducer characteristics into consideration, an improved signal model is proposed. Moreover, the corresponding sophisticated posterior distribution is derived and a similar sampling strategy is developed. This modification of signal model is more appropriate for the measured ultrasound echo signals and thus the resulting estimates for the model parameters can be used to characterise the scatterers in a more realistic way.

1.4 Overview of the thesis

The rest of this thesis contains a background chapter on ultrasound contrast imaging, which gives a review of recent developments on ultrasound imaging techniques and describes a procedure of ultrasound experimental measurements. It is then followed by five chapters of the major work. The conclusion is presented in the last chapter and the future research paths are also proposed.

Chapter 2 presents an evolution of imaging modalities from harmonic imaging to ultrasound contrast imaging techniques. In particular, a variety of investigations on pulsing excitation schemes to detect MBs in soft tissue. Moreover, the experimental setups for actual measurements of ultrasound scatterers are briefly introduced. They provide a basis of analysing measured ultrasound echo signals for the rest of this thesis.

Chapter 3 introduces several *non-parametric* estimation methods to characterise measured ultrasound echo signals in the time domain and the frequency domain separately. In the time domain, by combining the envelope detection method using Hilbert transform and voice activity detection (VAD) techniques, a new

method is proposed to estimate the temporal content of ultrasound echoes. In the frequency domain, estimation based on multitaper spectra is adopted as one of the most popular spectral estimation methods. Although the performance as their own is not satisfactory, they can be used as initial guesses for a specified model parameters in *parametric* estimation.

Chapter 4 proposes a signal model for ultrasound echo signals in a *parametric* Bayesian framework, according to the understanding of measured echo signals and the physical mechanism that generates them. It contains both temporal and spectral information of typical echo signals. Moreover, the proposed model is also compared to other models to demonstrate its appropriateness and advantages. Using the *parametric* estimation method, extracting the characteristics of ultrasound echo signals equals to estimating parameters of the proposed signal model. Based on Bayesian inference, the posterior distributions of the model parameters are obtained with the given priors and the likelihood function. As it is in a high dimension, the estimates of parameters cannot be achieved straightforwardly. Some popular sampling schemes are required.

Chapter 5 reviews approximating computations for Bayesian inference as the required sampling strategies, especially with the focus on Markov chain Monte Carlo (MCMC) algorithms. It covers from the very basic concepts to more advanced techniques in detail. They are used to sample from sophisticated posterior distributions of the model parameters. The *parametric* estimation methods developed in this thesis are all built on the basis of the described techniques and algorithms.

Chapter 6 develops a reversible jump MCMC (rjMCMC) algorithm to simulate samples from the sophisticated posterior distribution of parameters in the signal model proposed in Chapter 4. Then the model parameters can be estimated in a statistical way. Furthermore, the estimation results from Chapter 3 are used as initial values in the rjMCMC algorithm. They are updated iteratively to achieve refined estimates of the model parameters. In this way, an estimation system is created for analysing measured ultrasound echo signals. Moreover, the characteristics of ultrasound echoes are therefore extracted jointly rather than separately in both the time and the frequency domains.

Chapter 7 modifies the signal model proposed in Chapter 4 and develops a modi-

fied rjMCMC algorithm for the estimation of parameters in the modified model. The new model takes account of ultrasound transducer characteristics and regards them as part of the signal model parameters. Based on this new signal model, the posterior distribution of the model parameters is first derived. Then the corresponding rjMCMC algorithm, specially designed for the modified signal model, is used to simulate samples from the posterior distributions. In this way, the model parameters can be estimated and in particular, the true spectral content of echoes from ultrasound scatterers can be revealed.

Chapter 8 concludes the whole thesis and further research directions are suggested.

Chapter 2
Background knowledge on ultrasound contrast imaging

This background chapter first briefly reviews the evolution of the imaging modalities during the past decades. A variety of ultrasound imaging techniques are described and compared. The most encouraging technique is ultrasound contrast imaging since UCAs have been discovered and introduced to ultrasound imaging approaches. The image quality has thus improved to a large extent. Followed by discussion of the development in excitation pulsing schemes, the experimental setups for *in-vitro* measurements, which formulate the fundamental basis for analysing the measured ultrasound echo signals in later examinations, are presented in detail.

2.1 Evolution of imaging modalities

The ability to detect the presence of contrast MBs in blood or tissue is very important in enhancing the quality of ultrasound images. In the early days, MBs are used to enhance the backscattered signals with conventional imaging techniques. More specifically, when an ultrasound wave hits an MB, there will be two kinds of responses. One is from the interface between MBs and the other is from the surrounding tissue. They exhibit two different acoustic impedances. Under insonification of ultrasound, the size of an MB decreases in the positive cycle of the ultrasound wave and increases in the negative cycle. The volume pulsation of an MB is frequency dependent and shows a clear maximum at a specific frequency, which is often termed as a resonance frequency. In UCI, the volume pulsation results in an enhancement of the backscattered signal from MBs, which can be used to detect MBs in the presence of tissue.

Nevertheless, in hypo-echoic regions or very small blood vessels where the number of MBs is low, and the scattering strength is also low, echoes from MBs will be masked by echoes from the surrounding tissue. Therefore, the detectability of MBs is decreased. Nowadays, more and more new imaging techniques are developed to overcome this shortcoming. Most of them are based on specific properties of contrast

MBs and the main objective is to improve the SNR and the CTR of the received echo signals.

2.1.1 Fundamental imaging

Fundamental imaging mode is a simple and effective imaging technique [35]. In this mode, an increase of grey-scale value demonstrates the effectiveness of the employment of contrast MBs. Unfortunately, when the ratio of the blood volume to tissue is less than 10%, for example, for the myocardium, the increase of signal strength from MB responses will be very low and thus the fundamental grey-scale imaging will have poor detectability of contrast MBs. Nowadays, new imaging techniques are being developed according to more understandings of the interaction of MBs.

2.1.2 Harmonic imaging

Harmonic imaging is based on harmonic responses of ultrasound scatterers [36]. For lower amplitudes of an ultrasound wave, the relative expansion and compression of an MB is the same whereas for higher amplitudes, compression retards relative to expansion and thereby non-linearity occurs [35]. It means that the MBs get bigger and smaller in sympathy with the oscillations of pressure caused by the incident waveform. In particular, at their resonant frequency, they will absorb and scatter ultrasound with a high efficiency. Once echo signals are obtained, they are filtered with a band-pass filter to extract their corresponding second harmonics. As the echoes from contrast MBs have a relatively higher level of harmonics of reflection compared to tissue echoes, a much larger difference can be found when using the second harmonics rather than using the fundamentals. Therefore, the CTR at the second harmonics is substantially improved compared to the CTR at the fundamentals.

However, a traditional excitation pulse has a short duration in the time domain, which means it has a broad band for each frequency component in the frequency spectrum. The large broadband property will usually result in a significant overlap between the fundamental band and the second harmonic band in the frequency spectra, and will thus make it difficult to extract the harmonic components. Power Doppler imaging, which combines grey-scale image and Doppler image and display the strength of the Doppler signal in color, can help with it but is susceptible to interference from clutter caused by tissue motion [37]. A further limitation is that tissue may also have nonlinear propagation, which creates harmonics. Subharmonic imaging de-

tection may help to circumvent this problem, yet may cause the loss of the imaging resolution due to the narrow-band incident signals [38].

2.1.3 Multi-pulse schemes

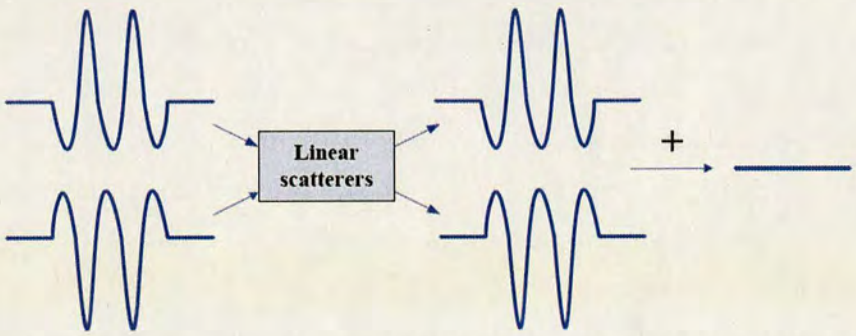
In order to overcome the limitations of the passive schemes based on single pulse excitation, in particular to overcome the increase of the trade-off between the CTR and imaging resolution, active multi-pulse schemes are proposed. There are two major techniques: pulse inversion [39] and power modulation [40]. The main advantage is that they can operate over the entire bandwidth of the received echo signals and thus achieve superior imaging resolution.

2.1.3.1 Pulse inversion

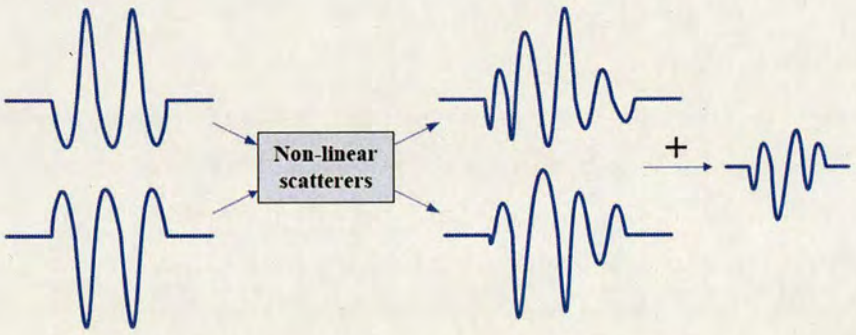
The basic operating principle of pulse inversion (PI) lies in the relative levels of even harmonics in received echoes. In the form of PI, a regular broadband ultrasound pulse and a phase inverted copy of this pulse are sent into the medium alternately. The received echoes are obtained by adding the resulting echoes from the phase inverted pulse to the echo from the not phase-inverted pulse. Figure 2.1 displays a simple illustration theoretically. If the scatterer is tissue, the result is a full cancelation of the responses and thus becomes zero in theory, as shown in Figure 2.1(a). If the scatterer is a contrast MB, the received echo will not be fully canceled and may produce some frequency components that are not present in the excitation. The rest of received echo is related to the degree of non-linearity as displayed in Figure 2.1(b). Unfortunately, in actual measurements, tissue will also have some non-linear propagation. This may cause mistakes in detection of the presence of MBs.

2.1.3.2 Power modulation

Furthermore, power modulation (PM), also denoted as amplitude modulation (AM), performs by sending two successive pulses that are equal in pulse shape but have different amplitudes. The first received echo is scaled and then subtracted from the second echo. The resulting echo will be canceled out in tissue whereas still has some remaining frequency components in contrast MBs. There are also simple illustrations of the principle of PM applied to both linear and non-linear ultrasound scatterers in Figure 2.2. In contrast to PI, only the fundamental frequency will be highly suppressed and both odd and even harmonics are remained if PM is used. The reason is that the



(a) Illustration of PI applied to linear scatterers.



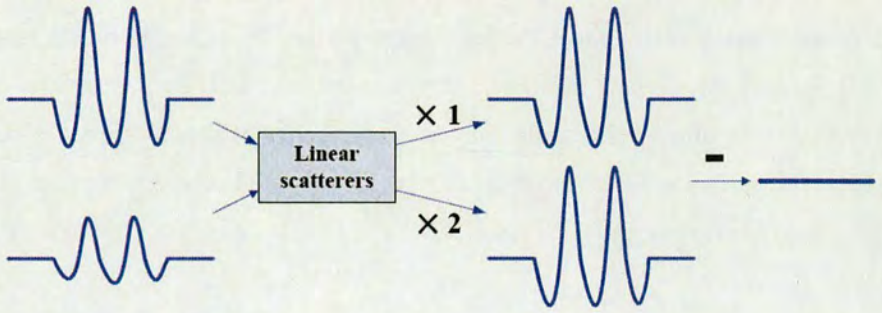
(b) Illustration of PI applied to non-linear scatterers.

Figure 2.1: *An illustration of PI applied to linear and non-linear scatterers in theory.*

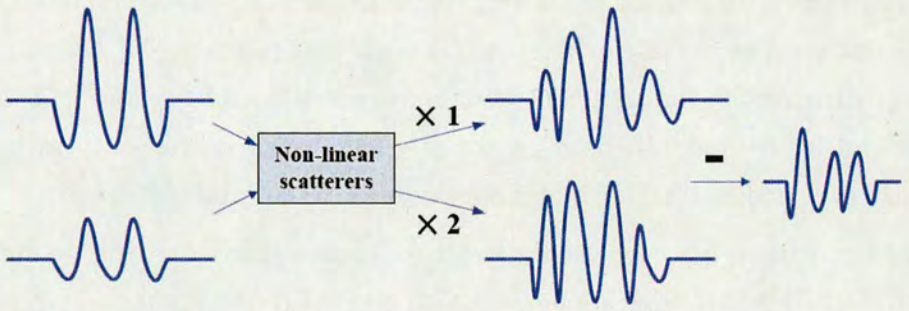
higher amplitude pulse will generate more harmonics than the lower amplitude pulse relative to the fundamental component and thereby, the harmonics will not be fully suppressed by the subtraction [19].

2.1.3.3 Multi-pulse imaging

Being performed on their own, PI gives a maximum amount of non-linearity at the even harmonics and excludes all the odd harmonics, and PM has the largest component at the second harmonic and only suppress the fundamental frequency component. The combination of PI and PM, denoted as pulse inverted amplitude modulation (PIAM), can eliminate their respective shortcomings [41]. Moreover, by introducing three or more pulses in the transmit sequence, the rejection of clutter due to the tissue motion will be increased and the SNR will be improved [42]. Michalakis [42] proposed various combinations of phase and amplitude modulations so as to isolate the funda-



(a) Illustration of PM applied to linear scatterers.



(b) Illustration of PM applied to non-linear scatterers.

Figure 2.2: *An illustration of PM applied to linear and non-linear scatterers in theory.*

mental from the harmonics. Nevertheless, the main disadvantages of the multi-pulse schemes are that artifacts will arise resulting from tissue motion and the frame rate will also be reduced.

2.1.4 Double frequency excitation

In 2004, Bouakaz et al. [12] proposed a new imaging technique using an incident pulse, which has an low frequency (LF) component, followed by an high frequency (HF) component. The LF component, denoted as a modulating signal, is used to modulate the size of an MB by inducing slow oscillations. It alters the size of the MB between the compression (positive cycle of the pressure) and expansion (negative cycle of the pressure) phases. The HF component, denoted as a detection signal, is used to image the MB. It will sense the MB at two stages: small and large size in accordance with the phase of the LF signal. It has been observed by a high-speed camera system in ultra-

sound experiments [12] that a larger response is obtained at the compression phase when compared to the response at the expansion phase. This change of MB responses is critically important. When non-oscillating scatterers, such as tissue, are present, the response will be almost the same during both phases of the LF signal. Therefore, this physical properties will help to increase the distinction between contrast MBs and tissue, which will improve the CTR finally.

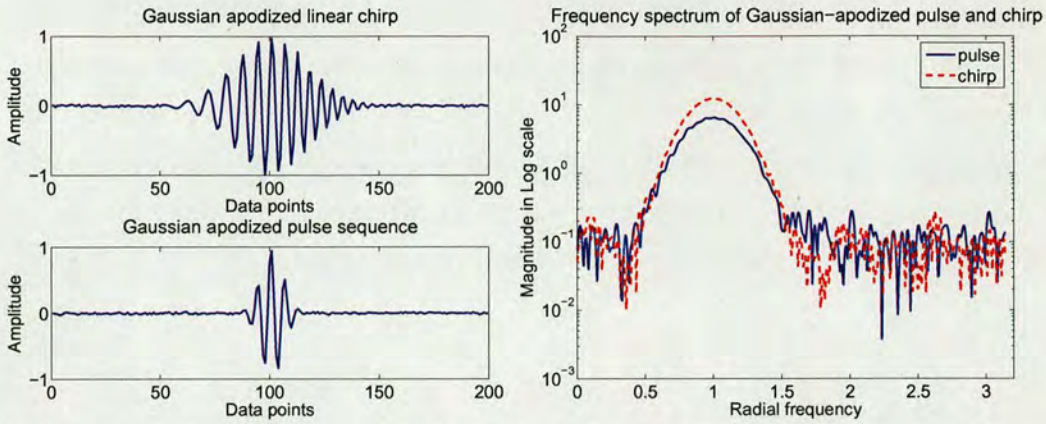
2.1.5 Coded excitations

In non-destructive contrast agent imaging, the maximum allowed MI limits the peak transmission pressure to avoid bubble cavitation and tissue damage [43]. Consequently, non-destructive contrast imaging will produce images with a relative low SNR. As the noise level is fixed by the system design, the only way to increase the SNR is increasing the signal level, which is limited by the bubble destruction threshold.

Coded excitations has been used in radar applications since the 1950s to improve the SNR [44]. They have been applied to UCI theoretically and through experimental evaluation by Borsboom [19]. They are employed to increase the signal energy by using longer pulses and a compression filter at the receiver instead of increasing the peak transmitted amplitude. Therefore, the SNR can be improved. Moreover, it has been demonstrated that the bandwidth and duration of the incident pulse will have influence on the generation of harmonics for MB responses whereas tissue harmonics generation mainly depends on the peak pressure of the incident pulse. Under this circumstance, the CTR can be increased.

Furthermore, from the perspective of axial resolution, longer incident pulses used in multi-pulse excitation schemes will generate larger oscillation amplitudes of the bubble wall and increase the generation of harmonics compared to shorter pulses with the same peak pressure, yet will compromise the axial resolution. Coded excitation, however, can combine longer pulses with good axial resolutions, which can increase the penetration depth, after some simple processing of the received signals.

There are two categories of codes that are suitable for ultrasound applications [45]. One is based on frequency modulation, such as chirp excitation; the other is based on phase modulation, such as Barker codes and Golay codes. They will be briefly introduced in Section 2.1.5.1 and Section 2.1.5.2 respectively.



(a) Gaussian apodized pulse sequence and linear chirp in the time domain.

(b) Gaussian apodized pulse sequence and linear chirp in the frequency domain.

Figure 2.3: *Pulse sequence and linear chirp apodized with Gaussian window in the time and the frequency domain.*

2.1.5.1 Chirp excitation

The most important code in frequency modulation is the linear chirp. It is a long sinusoidal burst with an instantaneous frequency changing linearly in time. Figure 2.3(a) shows examples of a sinusoidal pulse sequence and a linear chirp apodized with the same Gaussian window in the time domain. Their corresponding frequency spectra are displayed in Figure 2.3(b). It is shown that although two signals have similar magnitude spectra, the linear chirp has a longer duration of pulse than the pulse sequence.

In order to recover the axial/range resolution, the received echo signal needs to be processed using pulse compression techniques, which are mainly used in radar and sonar to augment the range resolution and the SNR [46, 47]. The compression techniques are achieved by modulating the transmit pulse and correlating it with the received pulse. Pulse compression is an example of matched filtering. A matched filter is usually implemented in telecommunications by correlating a known template, with an unknown signal to detect the template present in the unknown signal. It is often regarded as the optimal linear filter for maximising the SNR in the presence of additive white Gaussian noise. In radar signal processing, matched filters are widely used to investigate the reflected signal for common elements of the transmit signal.

Traditional compression acts on the same bandwidth as the excitation chirp and the employed matched filter has an impulse response equal to the time inverse of the chirp used as excitation. After the compression, the resulting signal has better axial

resolution than the initially received signal. Unfortunately, the harmonics from the received MB signal are suppressed by the conventional compression. Borsboom [19] proposed a new type of matched filter. Instead of using the same chirp for excitation and compression, the matched filter has double frequency at every point compared to the excitation chirp. In this way, the second harmonic from MB responses can be extracted by using the proposed matched filter in compression.

2.1.5.2 Other coded excitation

Apart from the chirp excitation, Barker and complementary Golay codes are another type of pulse compression technique using phase coding. They are able to alternate the phase of subsequent parts of the sinusoid over a fixed set of phase values [19]. The major advantage of Golay sequences is the zero-range side lobes [48]. Furthermore, the pulse inversion, combined with coded excitation, can also suppress the neighboring side lobes of harmonics.

However, coded excitation also suffers from the limitation of reduced frame rate. Moreover, the pulse compression ratio of the phase-coded compression is lower than in the chirp case and the compression is very sensitive to frequency changes due to the Doppler effect and thus the artifacts will arise from tissue motion.

2.1.6 Destruction based detection

The aforementioned techniques are all at a low MI, which can avoid the destruction of encapsulated MBs. However, when at a high MI (often MI > 0.1), the encapsulated MBs will be destroyed by releasing their gas content into the blood and thus become free MBs. The characteristics of reflections of encapsulated MBs and free MBs to the ultrasound will differ. As the free gas MBs have stronger scattering than the encapsulated ones, MBs can be detected by comparing the echo signals before and after the destruction. In this way, the CTR can be improved. Moreover, Borsboom [19] designed an experiment to compare the contrast MB destruction induced by pulse and chirp excitations. The measured results show that the intensity of MB destruction with chirps might be higher than that in the case of the destruction with pulsed waves. Therefore, chirps can provoke more changes in the response of MBs than pulsed waves.

2.1.7 Discussion

Since the introduction of contrast MBs into ultrasound imaging, a large amount of research has concentrated on the development of detecting MBs present in soft tissue. Although some of the presented detection methods perform well and are built into commercial ultrasound equipment, they are generally based on theoretical assumptions and not using much knowledge about the measured contrast MBs. In addition to the non-linear signature of the contrast MBs, no other characteristics are used in the current detection methods. Although intensive studies about the detection of MBs have been carried out in the past decades, there are still a lot space for improvement in this area, if some other specific knowledge of MBs can be found and incorporated.

This thesis steps back from detecting the high scattering strength and high non-linearity of contrast MBs, instead it explores more characteristics of ultrasound scattered signals and reveals true behaviour of MBs in soft tissue from a signal processing point of view. By developing specially tailored signal processing techniques, and proposing new algorithms specially designed for echo signals from ultrasound scatterers, the characteristics of reflections from contrast MBs and surrounding tissue can be exploited. In order to extract the characteristics of MBs and tissue from the very basic level, investigations are carried out based on experimental measurements using a simple six-cycle sinusoidal incident pulse with a single frequency component. The detailed experimental setup will be presented in Section 2.2.

Furthermore, unlike traditional ultrasound imaging techniques in which the waveform design and signal detection are two separate problems, a joint probabilistic framework in which the transmit pulse waveform is adapted to the measured MB characteristics, and the receiver adapted to the noise statistics and pulse transmit pulse characteristics, can be proposed if more understanding and knowledge of ultrasound scatterers can be learnt using some advances in signal processing techniques. It will explore a new research area of detecting contrast MBs and discrimination of them from soft tissue in the near future.

2.2 Experimental measurements for single ultrasound scatterers

As addressed in Section 1.2.2, the analysis of echo signals from single MBs is more useful than analysis of those from MB populations in revealing the measured MB

behaviour. In [2, 3], it was demonstrated in an experimental measurement system that there is considerable flexibility of control and ability to produce a large output of backscatter data from individual MBs. Their measurements are designed for single scatterers, especially for single MBs. Based on these measurements, some advanced signal processing techniques can be employed and developed to analyse the echo signals from single scatterers.

2.2.1 Contrast MBs and SCSs

The contrast MBs used in this study were Definity® (Bristol-Myers Squibb Inc, MA, USA), which were lipid coated and the gas was perfluorocarbon. The mean diameter of the MBs ranges from 1.1μm to 3.3μm. Moreover, the size of 98% of the microspheres are less than 10μm. The resonant frequency of oscillation of an MB in an ultrasound field depends on its size and is reversely proportional to square root of the third power of its radius. For example, for a 3μm diameter, the resonant frequency is about 3.3MHz.

The solid copper spheres (SCSs) (Goodfellows Ltd., Cambridge, UK), most of which have radii between 29.5μm and 57.5μm, were chosen in the experimental measurements due to two major reasons. First, they are rigid and cannot be compressed nor expanded, thus SCSs can be used to mimic the tissue behaviour. Second, SCSs can be regarded as reference materials of similar or larger size to contrast MBs in recovering true characteristics of MBs from measurements. Previous investigations used a plane surface, which is subjected to a range of acoustic pressures [15,30]. They are inappropriate for being regarded as references since contrast MBs are point scatterers. As the scattering from small SCSs has been well predicted by classical theory in [49,50], it can be combined with accurate measurements to calibrate the ultrasound receiver in the acoustic field.

2.2.2 Calibration of transmitted field

In experimental measurements described in [3], the transmitted field was calibrated using a membrane hydrophone. It is a Polyvinylidenfluorid (PVDF) piezoelectric membrane hydrophone (Precision Acoustics Ltd., Hampton Farm Business Park, Dorchester, UK), with an active area of 0.2mm, which is denoted as B in Figure 2.4. Moreover, the hydrophone has a submersible preamplifier that provides a gain of 8dB and 50Ohm output. This preamplifier is powered by a DC Coupler, which is denoted as

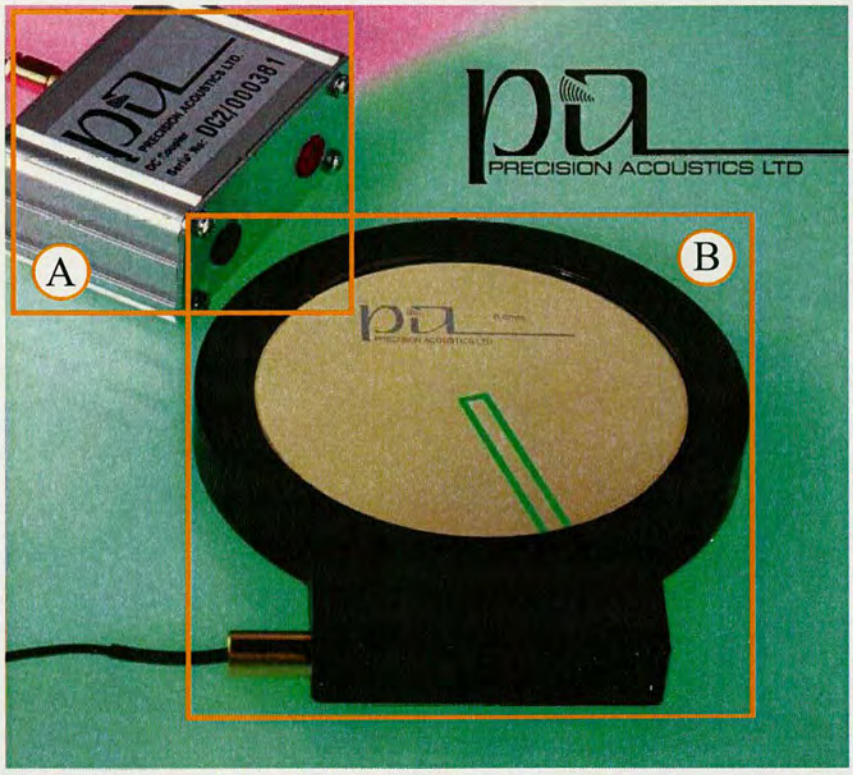


Figure 2.4: *The membrane hydrophone used in the experimental measurements [1].*

A in Figure 2.4. This membrane hydrophone is used to calibrate the incident pulse in the experiments. There are 21 different transmit pulses. They are all sinusoids with six periods. The only difference is that their fundamental frequency varied from 1.2MHz to 4.0MHz. The amplitude of the transmit pulse was controlled and set to 550kPa peak negative pressure, which indicates a low MI in the measurements.

Figure 2.5 illustrates how the transmitted field is calibrated using the membrane hydrophone and an ultrasound transducer. In this figure, the water tank is filled up with degassed water. Both the hydrophone and the transducer are immersed into the water and are put in parallel. The transducer is kept steady in the Z-axis and can be moved along both the X-axis and the Y-axis. The hydrophone is kept steady in X-axis and Y-axis whereas can be moved along the Z-axis. Moreover, the hydrophone can also be rotated so as to investigate the transmitted field in all possible directions and to achieve high precision of alignment between the transducer and the hydrophone. Along the Z-axis, the measurements are taken at a depth of 7.5cm where MBs and SCSs will be examined. At this depth, the transducer is moved along both the X-axis

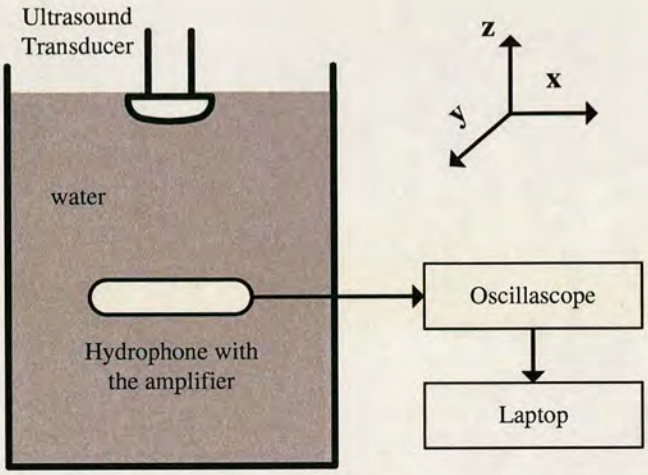
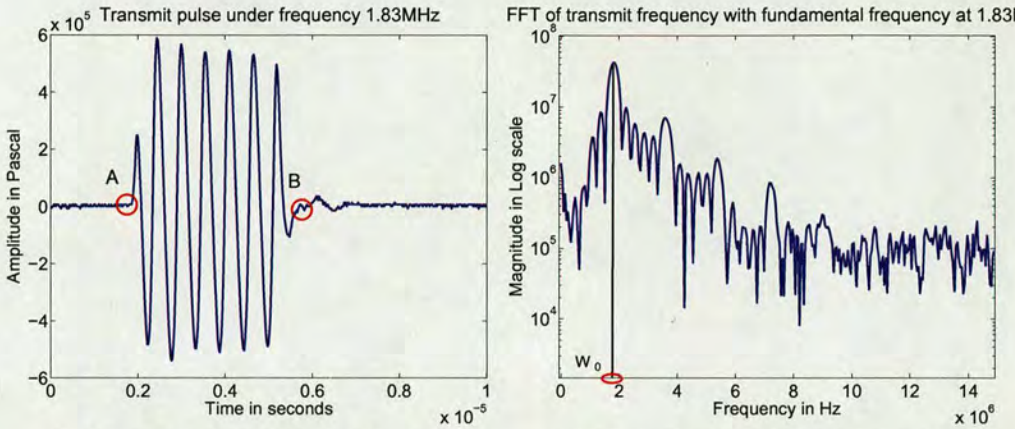


Figure 2.5: *An illustration of how the transmitted field is calibrated.*

and the Y-axis and the hydrophone is rotated until the signal intensity attains the maximum value, which indicates the centerline is reached. Furthermore, the transducer is connected to a computer through an AFLink USB interface. The hydrophone's preamplifier is connected to a DC coupler which is connected to an oscilloscope. Then the oscilloscope is connected to another computer under control. In this way, all the transmit pulses can be captured in the experiments.

Figure 2.6(a) displays an example of transmit pulses with the fundamental frequency at 1.83MHz in the time domain. The pulse duration is defined between the point A and B in the figure, which is a six-period sinusoidal pulse segment. The frequency spectrum of the defined pulse segment is shown in Figure 2.6(b). It can be inferred from the Figure 2.6 that the transmit pulse has a dominant frequency component around 1.83MHz. Other weak second and third harmonics are present because of the non-linear propagation of ultrasound in water. In the process of the propagation, the positive pressure of the wave forces the water molecules to come closer with each other and thereby increases the density whereas the negative pressure forces water molecules increase the distance between each other. This difference will introduce new frequency components during the propagation of ultrasound in water. Moreover, if the separation of the fundamental frequency and the harmonics is required, the bandwidth of the filter need to be selected carefully in order to avoid the overlap of frequencies. With the increase of the frequency of transmit pulse, the difference between the fundamental component and other harmonics will differ and therefore, the



(a) Transmit pulse in the time domain.

(b) Transmit pulse in the frequency domain.

Figure 2.6: *Transmit pulse in both the time domain and the frequency domain.*

bandwidth of the selected filter cannot be fixed. It should be adapted to the transmit frequency, which will make the filter design become more difficult.

2.2.3 Acquirement of echo signals from ultrasound scatterers

A modified ultrasound transducer (Sonos5500 Philips Medical Systems, Andover, MA, USA) was used to acquire echo signals from both SCSs and MBs. Figure 2.7 shows a view of the equipment used in the experimental measurements. The left part of the figure, denoted by A, shows the ultrasound transducer, which transmits an ultrasound wave and receives its reflection from scatterers. The sensitive range of receive frequency is set between 1.2MHz and 4.5MHz. Then the raw echo signals from the scatterers are preamplified, collected and stored in a computer. The right part of the figure, denoted by B, is a set of acquirement facility, which contains a water tank and some perspex tubes.

2.2.3.1 Experimental setup for acquiring SCSs

In order to clearly explain how the acquirement facility works, Figure 2.8 shows simple illustrations of experimental setups for acquiring echo signals from SCSs and MBs. As the density of an SCS is greater than water whereas the density of an MB is less than water, the equipment setups for them are different.

In Figure 2.8(a), a beaker was used as a basic tank to measure the scattering from SCSs. The diameter of the beaker bottom is 12cm and the height of it is 15cm. A 4cm diameter hole was drilled at the base of the tank and a 25μm thickness Mylar® film

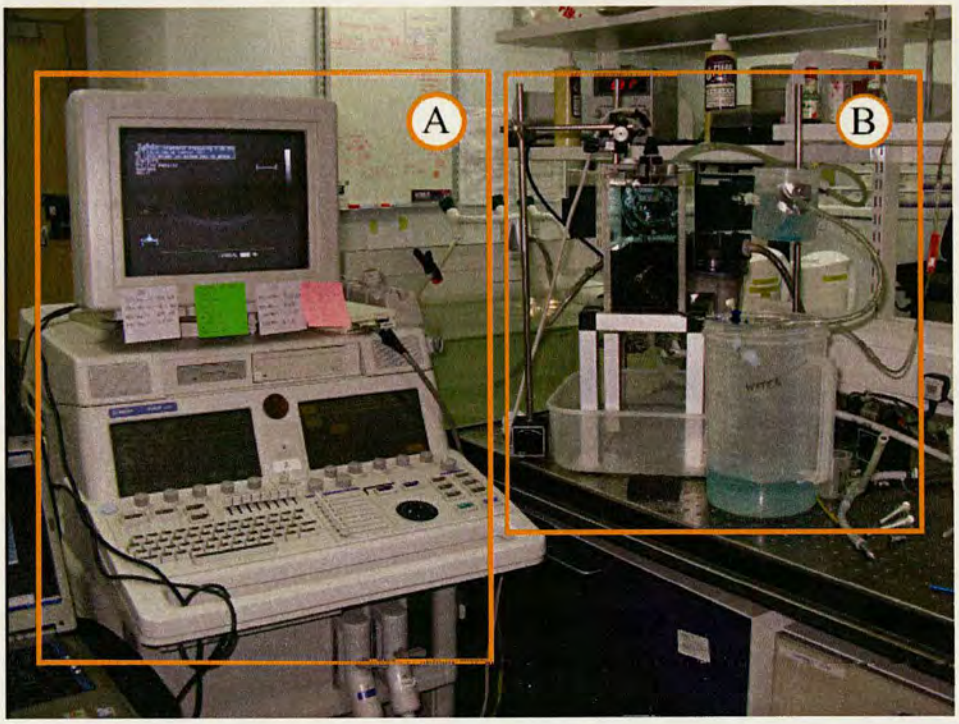
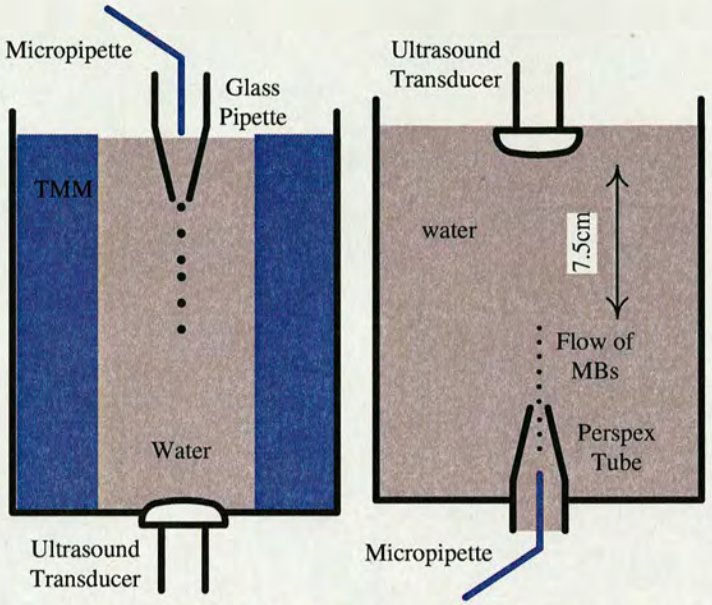


Figure 2.7: *The equipments used for acquiring ultrasound echo signals.*

was glued to the base to provide an acoustic window. The central cylindrical space defined by the circular acoustic window at the base were filled with degassed water. Other spaces are filled with the tissue-mimicing material (TMM), which has similar acoustic properties as tissue, such as acoustic velocity, attenuation and scattering coefficients and nonlinearity parameter. They are used because this setup can be compared to the actual environment in which MBs are surrounded by tissue. A glass pipette was placed at the top of the tank and its bottom tip with 1mm internal diameter was held at the center of the tank. Moreover, the measurements were carried out at 7.5cm from the bottom of the tank. The distance is chosen because it was able to put the SCSs under a near-plane wave and make positioning errors less critical [3]. Another micropipette was inserted through the larger glass pipette into the tank to ensure the alignment of the beam with the path of SCSs.

A small batch of SCSs with a variety of radii, ranging from 29.5μm to 57.5μm, were held inside the micropipette and then introduced to the tank individually. They descended towards the bottom of the tank with the guidance of the gravity. The ultrasound probe was positioned at the center bottom of the tank to achieve maximum



(a) Experimental setup for acquiring echo returns from SCSs.

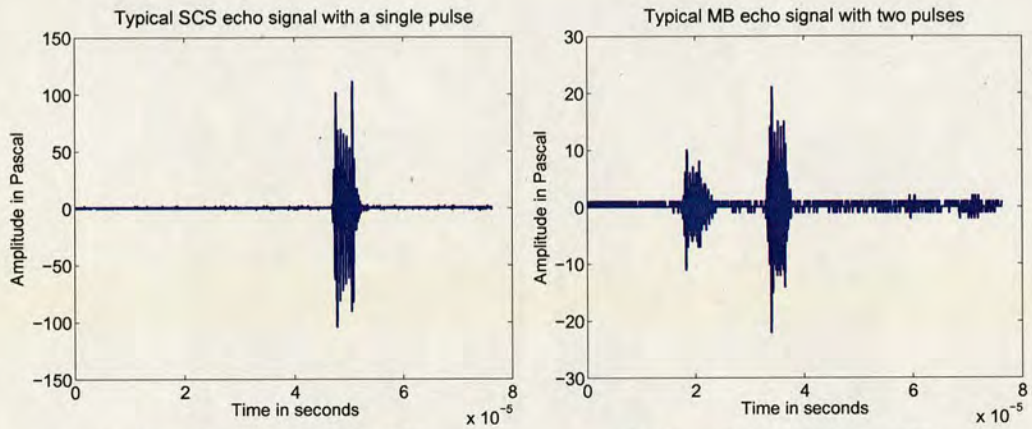
(b) Experimental setup for acquiring echo returns from MBs.

Figure 2.8: *Illustration of experimental setups for acquiring the echoes from SCSs and MBs.*

echo amplitude from the tip of the micropipette.

2.2.3.2 Experimental setup for acquiring MBs

The facility for acquiring echo signals from MBs is similar to the one used for SCSs. In Figure 2.8(b), the tank was filled with degassed water, which not produce echoes. Since the intensity of echo returns from MBs are low compared to those from SCSs, TMM will not be used in experiments. A Perspex tube was put at the center bottom of the tank with an 8mm internal diameter. The tip of the glass micropipette, whose diameter is approximately 100μm, was placed at the center of the Perspex tube. Single MBs are released from the micropipette individually. The rise time of the MBs is about 10 to 20 seconds, which is controlled by a stepper motor. The flow of them was directed towards the face of an ultrasound probe, which was placed at the top of the tank. The distance between the end of the tube and the face of the probe was 7.5cm, as shown in Figure 2.8(b). The tube and the micropipette were composed of a hydrodynamically focused flow system for isolation of single MBs in a well determined flow path. In this way, the insonation of single MBs can thereby be well calibrated. Moreover, the stream of MBs was first diluted by introducing filtered water into the infusion path. Then the



(a) The echo signal from an SCS.

(b) The echo signal from a MB.

Figure 2.9: *The echo signals from an SCS and an MB. (Note the highest signal strength in labels in two subfigures are different.)*

diluted suspension of MBs was supplied to the micropipette.

2.2.4 Discussion about the measured echo signals from SCSs and MBs

Once all the equipment has been set, echo signals from both SCSs and MBs can be obtained. According to the measurements, the echoes from SCSs are all single pulse signals whereas most of the echo signals from MBs have multiple pulses. Taking the transmit pulse displayed in Figure 2.6(a) as an example, the corresponding responses from a SCS and a MB are shown in Figure 2.9(a) and Figure 2.9(b) respectively. From the observation, there are quantised noise in the echo signals from MBs. In order to ease the mathematical calculation in the following analysis in the remaining thesis, the noise in both SCS responses and MB responses is assumed to be white Gaussian noise.

In this example, the SCS response has only one pulse segment and the MB response has two pulse segments. In Figure 2.9(b), the amplitudes of two pulses are different because the MBs may not in a perfect alignment and thus the ultrasound wave may not hit the center of MBs. Both of these echo signals have a total length of 1500 data points. The position of each pulse in echo signals has a relationship to the distance between the face of the ultrasound probe and the center of the scatterers in experimental measurements. Although these distances cannot be obtained directly, they can be calculated using a heuristic formula as follows:

$$distance = \frac{data\ points + offset}{scale}, \tag{2.1}$$

where the terms *offset* and *scale* are determined by the experimental setups for the acquiring of ultrasound scatterers. They are assigned as 608 and 255 respectively and the reason of choosing these values are discussed in [3] since the analysis of the ultrasound echo signals in this thesis is based on the measurements in [3]. The term, *data points*, denotes the center point of the detected pulse segment and thus is determined by the pulse location in the measured ultrasound echo signals. Moreover, the term, *distance* denotes the region of the most interests, which is set to 7.5cm in measurements in [3]. Furthermore, with respect to each incident pulse, there are 50 frames of SCS responses and 100 frames of MB responses with different pulse locations. These frames of responses are collected according to different distances between the ultrasound probe and the scatterers and thereby, signify various pulse locations.

As a result, in order to determine the *distance*, which is closest to 7.5cm, estimating the pulse locations in the echo signals in the time domain plays an important role in analysing the scattering from SCSs and MBs. Furthermore, as discussed in Section 1.1.2, contrast MBs have non-linear signatures in terms of the spectral contents of the corresponding transmit pulse, compared to tissue mimicking SCSs. This thesis will address the problems of estimating temporal and spectral contents of ultrasound echo signals, including linear scatterers, SCSs, and non-linear scatterers, MBs.

2.2.5 Transducer characteristics

During the process of acquiring echo signals from SCSs and MBs, the transducer is used as a receiver. However, the transducer has a band limited property, which will have influence on the estimation of spectral content of ultrasound echo signals. In order to remove this influence, the transducer needs to be calibrated and then can be used as prior information. The calibration is performed by comparing theoretical SCS echoes and experimentally measured SCS echoes. Moreover, both the magnitude response and the phase response should be taken into consideration. Detailed descriptions about the calibration procedure will be presented in Section 7.2. By incorporating the obtained transducer characteristics into the frequency estimation, the true spectral content of MB echoes can be recovered. Also, more characteristics of MBs can be revealed and thus more understanding of MB behaviour can be learnt.

2.3 Chapter summary

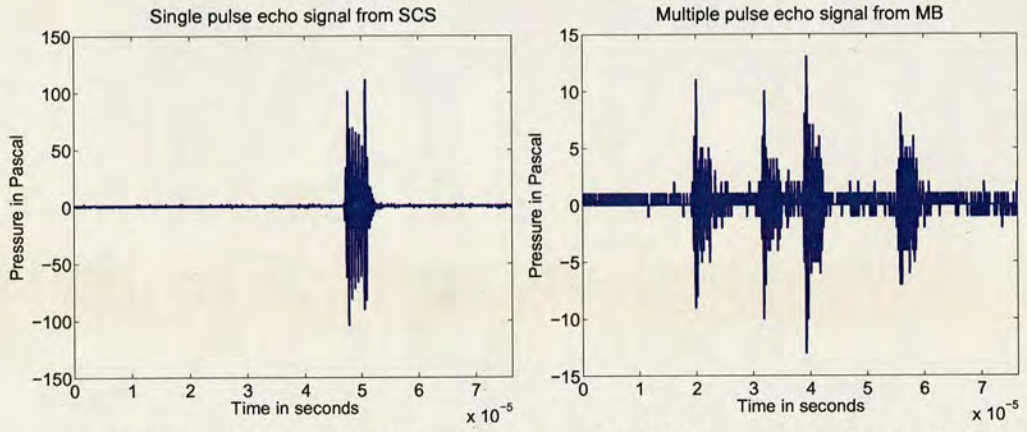
This chapter covers two parts of fundamental knowledge from the ultrasound imaging point of view in ultrasonics. First, the development of UCI techniques is reviewed from the very basic method, e.g. fundamental imaging and harmonic imaging, to more advanced approaches, e.g. coded excitation pulsing schemes and destruction-based detection. They provide the state-of-art in the research area of detecting MBs in tissue. Second, experimental measurements of single ultrasound scatterers, including MBs and SCSs, are described. All the subsequent analysis in this thesis are built on these measured ultrasound echo signals. According to the presented experimental setups and signal acquirements, signal processing techniques, specially designed for analysing ultrasound echo signals are required to characterise these scatterers. As a consequence, this chapter represents a stepping stone to the remainder of the thesis.

Chapter 3
Non-parametric estimation of temporal and spectral contents

A large number of signal processing techniques can be applied to the analysis of measured ultrasound echo signals in order to characterise the echo returns in the time domain and the frequency domain. These range from the discrete Fourier transform (DFT) to a full *parametric* model solved by Bayesian inference. This chapter considers several classical *non-parametric* techniques for temporal and spectral estimation. First of all, two popular signal processing methods for estimating pulse locations in signals, namely the Hilbert transform-wavelet denoising (HTWD) method and voice activity detection (VAD) algorithms, are introduced. Then a method, termed as HTWD-VAD, which combines these two techniques is proposed. Moreover, spectral estimation techniques based on Fourier analysis are utilised to estimate the spectral content of signals in the frequency domain. These techniques are first evaluated on a simulated signal and then are applied to the experimentally measured ultrasound echo signals.

3.1 Introduction

As aforementioned, the requirement of estimating temporal and spectral contents in ultrasound echo return signals based on experimental measurements originates from the lack of knowledge of real MB behavior and incomplete MB models in ultrasonics. If the temporal and spectral information of measured echo signals can be learnt with high accuracy, the research field of MBs will be broadened. Moreover, the estimation of SCS responses is also required to compare with MB responses. The experimental setup for acquiring the echo return signals from both SCSs and MBs has been illustrated in Figure 2.7 on page 28 and Figure 2.8 on page 29 and has been described in Section 2.2.3. As observed in the measurements in [2, 3], the echoes from SCSs only have a single pulse in the signal whereas the echoes from MBs usually have multiple pulses. Two typical examples of ultrasound responses are shown in Figure 3.1, one



(a) Single pulse SCS echo signal.

(b) Multiple pulse MB echo signal.

Figure 3.1: *Display of single pulse and multiple pulse echo signals from the experimental measurements in [2,3].*

is a SCS response with a single pulse, and the other is an MB response with multiple pulses.

In order to estimate pulse locations in the time domain and spectral content in the frequency domain for each pulse segment in ultrasound echo returns, a joint estimation system is proposed to characterise these echoes. It can be referred back to Figure 1.3 on page 9 in Section 1.3.1. This system consists of a *non-parametric* estimation part and a *parametric* estimation part. In this chapter, only the former part is considered. Figure 3.2 shows an extract from Figure 1.3.

Generally speaking, *non-parametric* methods make fewer assumptions about general data property, thus are more widely used than *parametric* methods, especially in situations when little knowledge about the data is available. Moreover, *non-parametric* methods are simpler and easier to use in many real world applications. By contrast, *parametric* methods use extra assumptions compared to *non-parametric* methods. They usually assume data come from a specific probability distribution and make inferences about the parameters of the distribution. If these assumptions are correct, *parametric* methods can produce more precise estimates. The challenge lies in how closely the mathematical model matches the actual physical process that produced the data. If a model is not well-suited to the data, the application of the model will become meaningless [51]. In this chapter, only *non-parametric* estimation methods are presented.

In subsequent sections, *non-parametric* methods are carried out respectively in the time and the frequency domains to estimate temporal and spectral contents for ultra-

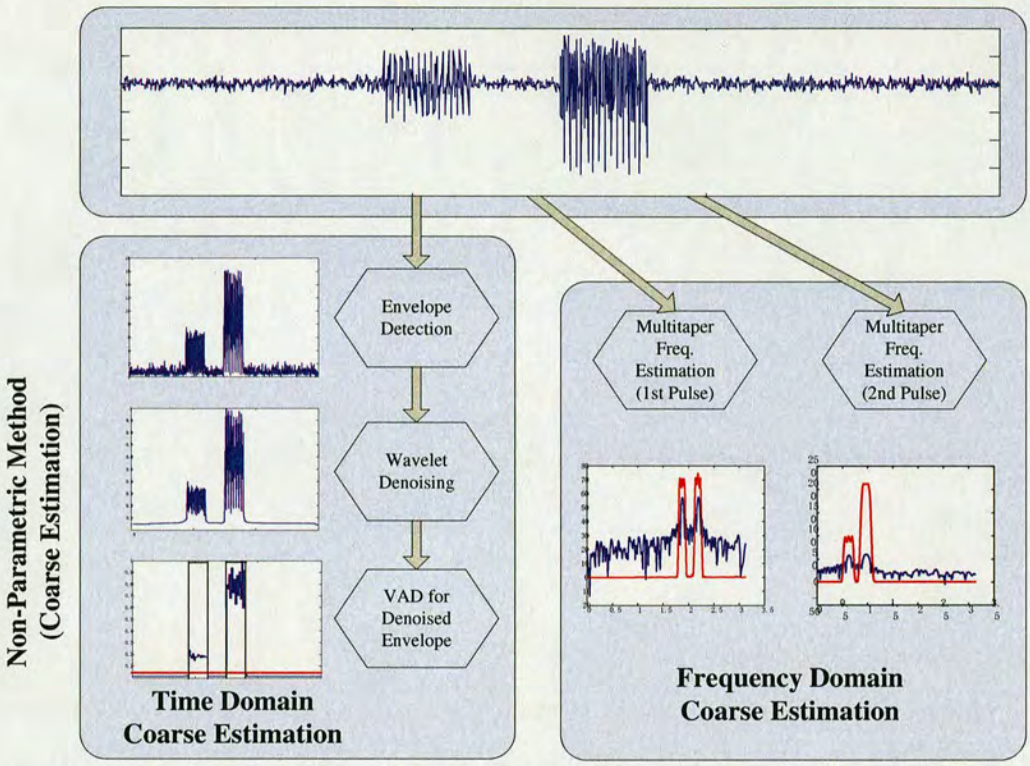


Figure 3.2: *Diagram of the coarse estimation part of the system.*

sound echoes. In the time domain, improved estimates can be achieved when combining the envelope detection technique and VAD algorithms, namely the HTWD-VAD method. This will be described in detail in Section 3.2.4. In the frequency domain, a low-variance spectral estimator, i.e. multitaper power spectrum, is computed to estimate the frequency components of a signal, which will be presented in Section 3.3.3. In Section 3.3.4, the STFT is adopted to provide estimates for multiple pulse signals, which have a time-varying property. Moreover, encouraging results and limitations will also be addressed.

3.2 Estimation of pulse locations in the time domain

Since some measured ultrasound echo signals have a single pulse whereas some have multiple pulses, estimation techniques for pulse locations must be able to detect both single pulse echoes and multiple pulse echoes without any prior knowledge. From this perspective, the HTWD method presented in Section 3.2.1 and Section 3.2.2, and the VAD algorithm introduced in Section 3.2.3, can both be used for estimation of pulse

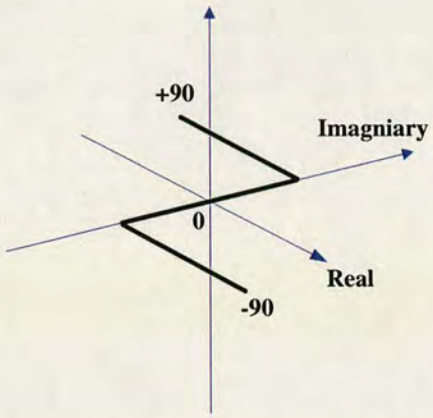


Figure 3.3: *A sketch diagram of the effect of the Hilbert transform on phase shift.*

locations.

3.2.1 Envelope detection

The Hilbert transform (HT) is a commonly used tool for envelope detection in various applications, especially in signal processing for amplitude modulated (AM) signals in tele-communications [52]. It can be implemented by applying the Fourier transform (FT) first and then multiplied by a factor, $(-j\,\mathrm{sgn}(\omega))$, in the frequency domain. The AM signals are often used for transmitting information via a carrier wave, usually a sinusoid. The energy of the transmitted signal varies according to the information that needs to be sent. In order to demodulate the modulated signal, the analytic representation of a signal based on the HT can be used. In this way, the envelope, indicating the original signal, can be extracted.

In principle, the HT only affects the phase response of a signal and has no effect on magnitude response at all, which can be described as: all negative frequencies of a signal get $+90$ degree phase shifts and all positive frequencies get -90 degree phase shifts. Figure 3.3 displays a sketch diagram of the HT on how it shifts the phase in the frequency domain. Moreover, a mathematical way of expressing the HT can be defined as follows with a given signal $g(t)$:

$$\hat{g}(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{g(\tau)}{t-\tau}d\tau = \frac{1}{\pi t}\otimes g(t), \tag{3.1}$$

where $\otimes$ denotes the convolution operation. More detailed information about the HT can be found in [53].

An analytical signal, which is composed of a real signal and its corresponding HT as the quadrature component, has a spectrum that only exists in the positive frequency domain. This analytical signal is often used to extract the envelope that contains the energy of an AM signal [53, 54]. For example, if the original AM signal is denoted as $x(t)$, then its corresponding analytic signal, which has no negative frequency components, can be expressed as: $x_a(t) = x(t) + j\hat{x}(t) = A(t)e^{j\Phi(t)}$, where $\hat{x}(t)$ is the HT of $x(t)$. $A(t)$ and $\Phi(t)$ represent the *amplitude envelope* and the *instantaneous phase* of the signal $x_a(t)$ as follows:

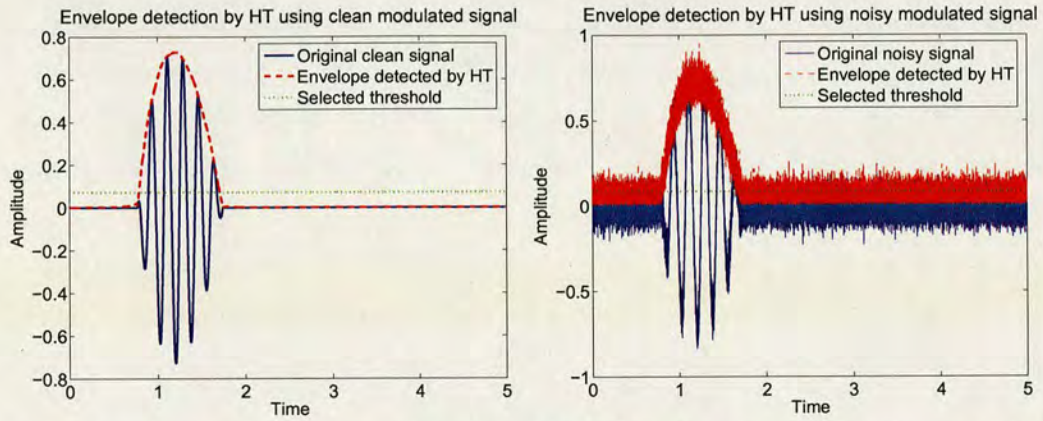
$$A(t) = |x_a(t)| = \sqrt{x^2(t) + \hat{x}^2(t)}, \tag{3.2a}$$

$$\Phi(t) = \arg\left[x_a(t)\right]. \tag{3.2b}$$

The envelope of the original signal $x(t)$ is the magnitude of the analytical signal $x_a(t)$, which can be used as an energy detector. Figure 3.4(a) displays an AM signal in the form of $x(t) = a(t) \times s(t)$ with no noise. The real baseband signal $s(t) = -t^2 - (3.4/t) + 5$ is quadratic. The reason of choosing it is that it has a similar shape as a typical ultrasound echo signal. Moreover, $a(t) = \sin(2\pi\omega t)$ is a carrier signal with a single frequency at $\omega = 5.6$.

In Figure 3.4(a), an original modulated signal curve, $x(t)$, is denoted in a blue solid line. The objective is to extract $s(t)$ from $x(t)$ using the HT. By calculating the magnitude of the analytical signal, the envelope of the original signal, $s(t)$, can be detected, which is denoted in a red dashed line. Once the envelope is obtained, a threshold must be selected carefully to determine the exact pulse locations. In a noise-free environment, as shown in Figure 3.4(a), the threshold is the only parameter that will have a large influence on the estimated pulse location. The thresholds can be chosen in proportion to the maximum value of the signal amplitude, ranging from 0.01 to 0.5. The green dotted horizontal line in Figure 3.4(a) denotes a selected threshold at 0.1.

In the presence of noise, the SNR becomes the dominant influence factor in estimating the pulse location. Since the signal $x(t)$ is non-stationary, the SNR of $x(t)$ becomes



(a) Envelope detection for the clean modulated signal.

(b) Envelope detection for the noisy modulated signal.

Figure 3.4: *Envelope detection for the clean and noisy signal using Hilbert transform.*

time-varying. In order avoid confusion, the SNR[1] is defined as follows:

$$\begin{aligned} \text{SNR(dB)} &\triangleq 10\log_{10}\frac{(\prod_{j=1}^{k_i} P_{s_j})^{\frac{1}{k_i}}}{(\prod_{j=1}^{k_i} P_{n_j})^{\frac{1}{k_i}}} \\ &= \frac{1}{k_i}\sum_{j=1}^{k_i}\left[10\log_{10}\left(\frac{P_{s_j}}{P_{n_j}}\right)\right], \end{aligned} \tag{3.3}$$

where $(\prod_{j=1}^{k_i} P_{s_j})^{\frac{1}{k_i}}$ is the geometric subband mean and $(\prod_{j=1}^{k_i} P_{n_j})^{\frac{1}{k_i}} = (P^{k_i})^{\frac{1}{k_i}} = P$ is the arithmetic noise mean. Moreover, k_i denotes the number of sinusoids in a single pulse segment. Each subband is defined within a very small bandwidth, $\triangle B$, with a center frequency at each frequency component, ω_{k_i}. The noise power can be represented by: $P = \sigma^2 \times \triangle B$, in which σ^2 is the noise variance. This definition of SNR emphasizes the contribution of each frequency component in a signal and is used as a tradition in frequency estimation papers, e.g. [33,34].

Adding a Gaussian noise to the clean signal shown in Figure 3.4(a) will create a noisy signal. The SNR is set to 20dB following the definition in (3.3). Figure 3.4(b) shows this noisy signal. The threshold chosen at 0.1 is denoted by the green dotted horizontal line in the figure. Nevertheless, it is very hard to determine the start and end points of the pulse with the selected threshold. The estimation of the start and end point will have a strong influence on the estimation of frequency components.

[1]This definition of SNR for a single pulse signal is used throughout the thesis.

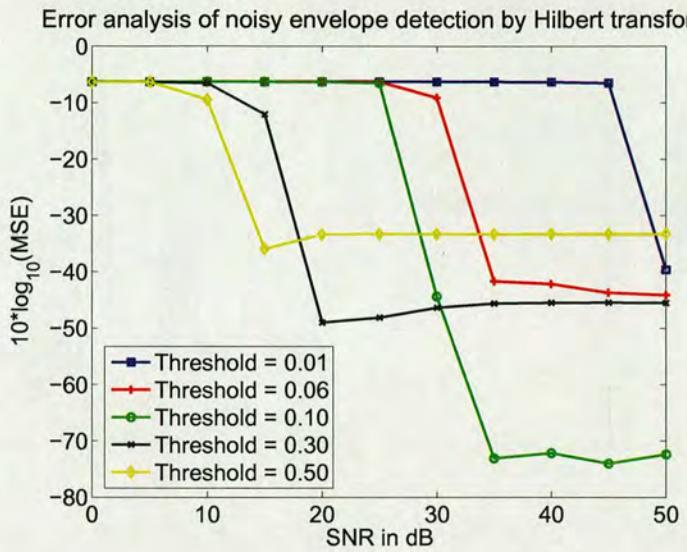
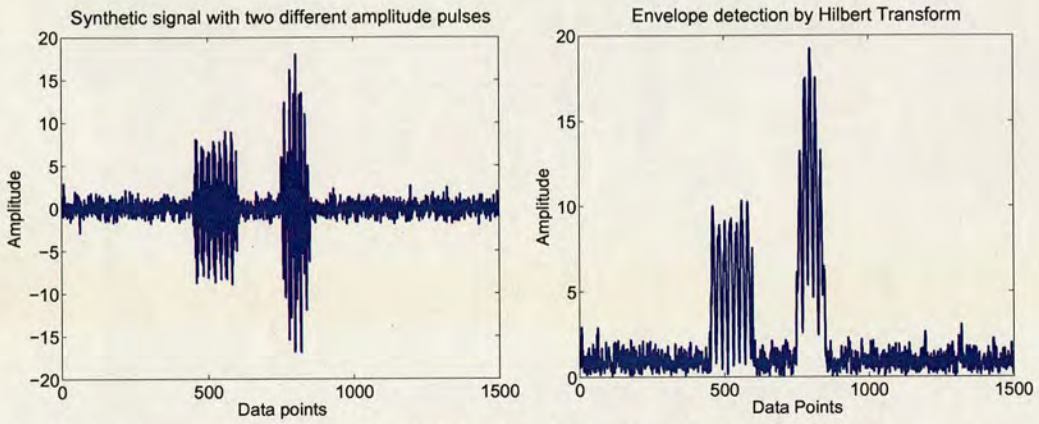


Figure 3.5: *Performance evaluation of envelope detection using Hilbert transform.*

Figure 3.5 shows the selection of threshold from 0.01 to 0.5 with different SNRs. The estimation performance is evaluated using mean squared error (MSE), in which the errors are calculated as the ratios of the difference between the estimated frequency values and the corresponding ground truths, to the full signal length. If the threshold chosen varies from 0.01 to 0.10, the errors will increase and the estimation can only be made possible when the SNR is higher than 50dB; if the threshold chosen varies from 0.10 to 0.50, the errors will also increase whereas the estimation can be made possible when the SNR is as low as 15dB. Therefore, when the SNR is lower than 15dB, the threshold at 0.50 is the most suitable choice whereas when the SNR is higher than 30dB, 0.10 is the most preferable threshold.

This HT based envelope detection approach is originally used to detect a baseband signal $s(t)$, i.e. the envelope of a modulated signal $x(t)$ in telecommunications. Nevertheless, if a signal is not modulated, the HT based envelope detection can still be used although the performance is not very satisfactory.

In order to evaluate the HT based envelope detection method on non-modulated signals, another simulated signal, which is not modulated, is employed. Figure 3.6(a) displays an example, which is analogous to the experimentally measured multiple pulse echo signal from ultrasound scatterers. Since this two-pulse signal will be used in later Chapters to give a comparison, details about how it is synthesized with various parameters can be found in Appendix A.



(a) Original two pulse signal embedded in the noise.

(b) Envelope detection by the Hilbert transform.

Figure 3.6: *Envelope detection for noisy signal by the Hilbert transform.*

As shown in Figure 3.6(a), there are two pulse segments in the signal and the SNR for this multiple pulse signal is time-varying. Therefore, a definition of the averaged SNR for multiple pulse signal is necessary. Assume there are m pulse segments, each segment i has its own SNR: SNR_i. Then the averaged SNR of the whole signal, $\overline{\text{SNR}}$, [2] is defined as the arithmetic average of all SNRs in all pulses, which is expressed in (3.4). Moreover, each individual SNR in a single pulse can be calculated according to the definition in (3.3).

$$\overline{\text{SNR}} \triangleq \sum_{i=1}^{m} \text{SNR}_i. \tag{3.4}$$

Figure 3.6(b) shows the detected envelope of the non-modulated signal in Figure 3.6(a) using the HT. Unfortunately, it is so noisy that an appropriate threshold is difficult to choose and thereby the start and end points of the envelope are difficult to determine. Accordingly, denoising of the envelope is necessary so as to achieve the estimated pulse locations of the synthetic signal.

3.2.2 Wavelet denoising techniques

Even though the envelope of a signal has been detected, determining the start and end points of multiple pulses in the signal is difficult if the signal is embedded in noise, especially with varying pulse amplitudes. Various techniques have been proposed in the literature for removing the noise from noisy signals.

[2] This definition of SNR for a multiple pulse signal is used throughout the thesis.

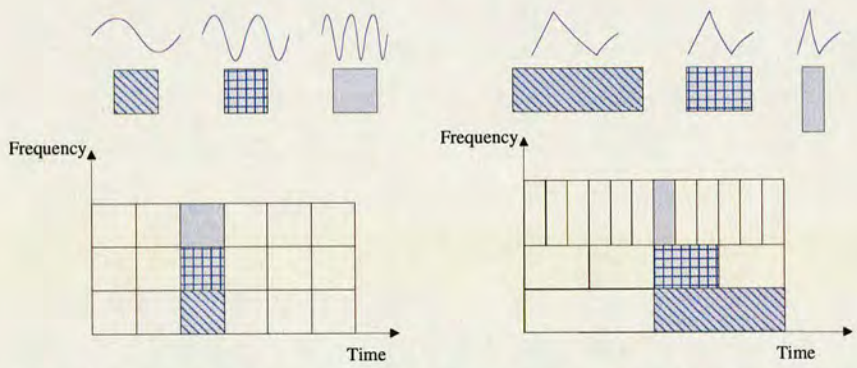
3.2.2.1 Denoising methods

The simplest denoising method is DFT-based denoising [55, 56]. In the DFT-based signal denoising, the DFT of a signal is taken. Then the transformed signal in the high frequency region where most of the noise is present, is attenuated or completely set to zero. Finally, the signal is reconstructed using an inverse DFT (IDFT). Using this DFT based approach, the high frequency noise can be removed.

Low-pass filtering approach is also a traditional method to remove noise [57]. It is able to smooth a signal by getting rid of the frequencies higher than a cutoff frequency. Several commonly used linear filters, such as Butterworth, Chebyshev and Elliptic filters, are compared in [57] with the consideration of different cutoff frequencies, pass-bands and stop-bands. However, if there are low frequency noise components present, this approach is not a good choice and the denoising performance is not effective.

Another choice for denoising is based on the STFT [58]. It applies a finite-length window, e.g. a Hanning window, to a specified block of signal and then takes the DFT of it. Therefore, a time-frequency representation of the signal, i.e., the STFT of the signal, can be obtained. This denoising scheme involves calculating the STFT of a signal, multiplying it by a 2D array mask, and reconstructing the signal from the modified STFT [59]. This 2D array mask is composed of ones and zeros. Ones represent the desired information that need to be preserved whereas zeros represent the noises that need to be removed. In practical applications, the zeros can be replaced by a certain small enough value and the ones can be replaced by a relatively large value. Nevertheless, if the window length is not properly selected, the performance of denoising will not be satisfactory.

Alternatively, matched filters [60] and notch filters [61] are usually utilised to remove noise. They are performed in the frequency domain and are tuned specifically to the frequency of the signal. The cancelation of noise after the filtering is thereby achieved. Nevertheless, the precise frequency of the noise needs to be known and fixed. Furthermore, the least mean squares (LMS) filter and the recursive least squares (RLS) filter, which are two common filters with adaptive and iterative gradient search algorithms, are introduced in [62, 63]. In the denoising process, the tap weights of filters are adapted by means of the LMS or the RLS algorithm. After a number of iterations, the noise in the signal will be diminished. However, the errors may arise when applied to realistic noisy signals.



(a) The diagram of time-frequency representation of the STFT.

(b) The diagram of time-frequency representation of the WT.

Figure 3.7: *The diagrams of time-frequency representation of the STFT and the WT.*

There is also attempt at using empirical mode decomposition (EMD) as a denoising tool [64, 65]. The EMD decomposes a noisy-signal into a number of amplitude and frequency modulated zero-mean signals, i.e., various intrinsic mode functions (IMFs). The energy difference between a noise-only IMF and the corresponding noisy-signal IMFs represents the signal information. In a denoising scenario, the IMFs containing signal information are used to reconstruct the denoised signal. In practice, however, the noise-only IMF can not be easily estimated based on the actual noisy-signal.

3.2.2.2 Wavelet denoising

In the past decade, denoising based on the wavelet transform (WT) becomes more and more popular in a number of applications. A detailed introduction to wavelets is given in [66, 67]. Wavelets are a mathematical tool to decompose a signal into different components and offer the ability to represent different levels of details present in the signal. Compared to the STFT for a time-varying signal, the WT is interpreted as a mapping of the signal in the time domain into a 2D function of both time and frequency. The WT uses variable size time windows for different frequency bands whereas the STFT uses the same size time windows for all frequency bands, as shown in Figure 3.7.

In the continuous wavelet transform (CWT), the energy of a signal, $y(t)$, is projected on different levels of continuous frequency bands or subspaces, which are scaled versions of a subspace at level 1. The level 1 of a subspace is denoted by a *mother*

wavelet $\phi(t)$. The *mother wavelet* is scaled by a and shifted by b to give a *child wavelet*:

$$\phi_{a,b}(t) = \frac{1}{\sqrt{a}}\phi\left(\frac{t-b}{a}\right), \tag{3.5}$$

where $a \in \mathbb{R}_+$ and $b \in \mathbb{R}$. Therefore, the projection of $y(t)$ onto the subspace of scale a has the form:

$$y_a(t) = \int_R \{WT_\phi y(a,b)\}\, \phi_{a,b}(t) db, \tag{3.6}$$

and the wavelet coefficients at the subspace of scale a can be expressed as:

$$WT_\phi y(a,b) =< y, \phi_{a,b} >= \int_R y(t)\phi^*_{a,b}(t)dt, \tag{3.7}$$

where $< \cdot >$ represents the inner product operation and * represents the complex conjugate operation.

In the discrete wavelet transform (DWT), a pair (a,b) defines a point in the right half-plane $\mathbb{R}_+ \times \mathbb{R}$ and it is sufficient to select a discrete subset of the upper half-plane to reconstruct a signal from the corresponding wavelet coefficients with integers m and n:

$$y(t) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} < y, \phi_{m,n} > \phi_{m,n}(t), \tag{3.8}$$

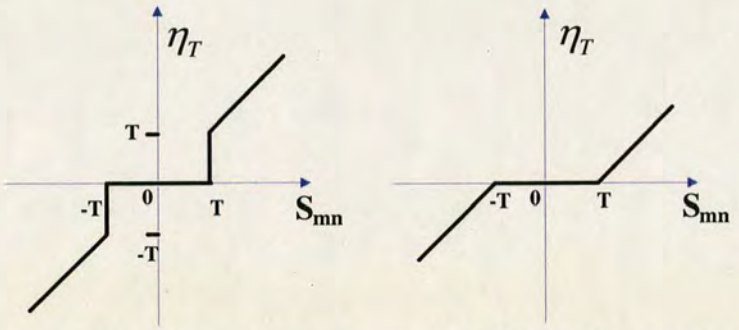
where $S_{mn} =< y, \phi_{m,n} >$ denotes the wavelet coefficients, and $\phi_{m,n}(t)$ are the corresponding *child wavelets*, given as:

$$\phi_{m,n}(t) = a^{-m/2}\phi(a^{-m}t - nb), \quad m,n \in \mathbb{Z}^2. \tag{3.9}$$

Since typical values can be assigned to $a = 2$ and $b = 1$ [66,68], the child wavelet basis in (3.9) will be simplified as:

$$\phi_{m,n}(t) = 2^{-m/2}\phi(2^{-m}t - n). \tag{3.10}$$

Furthermore, Mallat's multiresolution analysis (MRA) [68] is a tool for a constructive description of different wavelet bases in a more general framework, and is a design method for most of the practically relevant DWTs. It decomposes a signal into several scale levels, which are averaged and differentiated recursively. The resulting *detail* coefficients are neglected and the *averaging/approximation* coefficients are used to reconstruct the signal.



(a) Hard thresholding function. (b) Soft thresholding function.

Figure 3.8: *Illustrations of hard and soft thresholding functions.*

There are lots of applications which use wavelet decompositions, such as computer and human vision, data compression and denoising noisy data [69]. The application of denoising is the main target in this section. The signal denoising can be implemented based on the idea of thresholding wavelet coefficients S_{mn}, which is a way of throwing away unimportant details as noise. If the details are small, they can be omitted without affecting main features of the whole signal.

In general, there are two steps in the process of thresholding the wavelet coefficients when denoising signals. The first step is the choice of a threshold function. Two standard choices are hard thresholding and soft thresholding [70]. The functions are defined in (3.11a) and (3.11b) respectively, and are illustrated in Figure 3.8 with a given threshold T. These two thresholding functions are proved to give the best spatial adaptation in [71].

$$\text{Hard thresholding:} \quad \eta_T = \begin{cases} S_{mn} & \text{if } |S_{mn}| > T \\ 0 & \text{otherwise} \end{cases} \tag{3.11a}$$

$$\text{Soft thresholding:} \quad \eta_T = \begin{cases} \operatorname{sgn}(S_{mn}) \cdot (|S_{mn}| - T) & \text{if } |S_{mn}| > T \\ 0 & \text{otherwise} \end{cases} \tag{3.11b}$$

The second step is the choice of a threshold T. Two basic rules are used for selecting the threshold: a universal threshold, which depends on the length of a signal [72], and a threshold based on Stein's unbiased risk estimator (SURE) [71]. Alternative choice of selecting a threshold is minimax thresholding, first proposed for real signals in [71]. Then it is extended to the selection of a threshold for complex signals in [73]. The minimax thresholding rules are more conservative and more convenient if small

details are present in the noise range, compared with two basic thresholding rules [71].

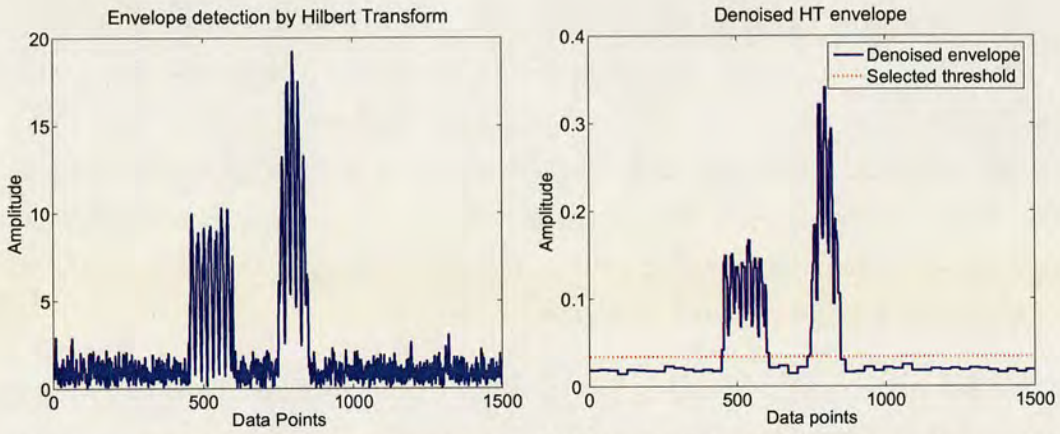
This technique is a breakthrough in dealing with noisy data because the denoising is performed without smoothing out the sharp structures. The result preserves a cleaner signal and shows important details as well. As a result, considering the necessity of minimization of estimation errors and the processing time to perform denoising, the wavelet denoising scheme by thresholding can be applied to denoising the HT envelope. The procedure is summarized as follows:

1. Decomposition: Choose Haar wavelet, which is the simplest and oldest wavelet when implementing the denoising [66]. Since the denoising is applied to the detected square-shaped envelope of a signal rather than a signal itself, Haar wavelet is selected due to its tendency to create square structures and thus is more appropriate for denoising the envelope. Moreover, choose level 5, which is a good trade-off between the denoising performance and the computational complexity. Therefore, compute wavelet decomposition of a noisy signal at level 5 using Haar wavelet.

2. *Detail* coefficients thresholding: For each level from 1 to 5, select a threshold and apply soft thresholding to the *detail* coefficients according to (3.11b). Then choose minimax thresholding as the selection rule of thresholding.

3. Reconstruction: Compute wavelet reconstruction based on the original *approximation* coefficients of level 5 and the modified *detail* coefficients from level 1 to level 5.

Take the simulated signal described in Appendix A as an example. In Section 3.2.1, Figure 3.6(a) on page 40 displays the original signal in the time domain and Figure 3.6(b) shows the detected noisy envelope using the HT. In this section, wavelet denoising technique is applied to the noisy envelope and the resulting denoised envelope is displayed in Figure 3.9(b). The reason is that a clearer envelope rather than a clearer signal is required in order to determine the start and end points in the signal. Figure 3.9(a) is shown to compare the envelopes before and after the wavelet denoising. As seen from the denoised envelope in Figure 3.9(b), the ripples at the bottom are almost removed by denoising. By contrast, the ripples at the top of the detected envelope still remain. This is because the wavelet denoising may regard the ripples as detailed structure characteristics of the envelope. Nevertheless, it is much easier to



(a) Envelope detection by the Hilbert transform.

(b) Normalised denoised envelope using wavelet denoising.

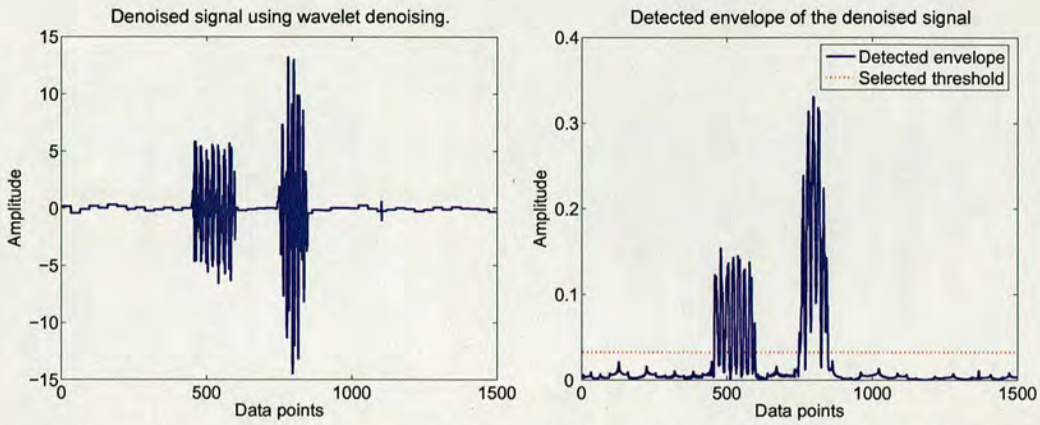
Figure 3.9: *Detected envelopes before and after the denoising process.*

select a proper threshold for the denoised envelope and thereby to determine the pulse locations.

Alternatively, the signal can be denoised first and then the HT is applied to the denoised signal. Figure 3.10(a) shows the denoised signal and Figure 3.10(b) displays the detected envelope of the denoised signal using the HT. Compare Figure 3.9(b) and Figure 3.10(b), the result of the former obtained by using the HT to detect the envelope first and then applying the wavelet denoising is termed as the HTWD method. The result of the latter achieved by wavelet denoising of the original signal first, then applying the HT to detect the envelope is termed as the WDHT method. According to the obtained envelopes, the HTWD method outperforms the WDHT method as there are still many ripples around the detected envelope in Figure 3.10(b), which still makes it difficult to determine the start and end points of the pulses in the signal. Moreover, the ripples at the bottom in Figure 3.9(b) are almost flat whereas Figure 3.10(b) shows several small spikes, which will also have an influence on choosing a threshold.

3.2.3 VAD algorithms

In Section 3.2.1, the envelope detection technique based on the HT is introduced to estimate multiple pulse locations in measured ultrasound echo returns. Nevertheless, the detected envelope is so noisy that other denoising techniques are required. Therefore, the wavelet denoising is introduced in Section 3.2.2 to remove the noise and determine the edges of each pulse. In this section, an alternative method is adopted to estimate



(a) Denoised signal using wavelet denoising.

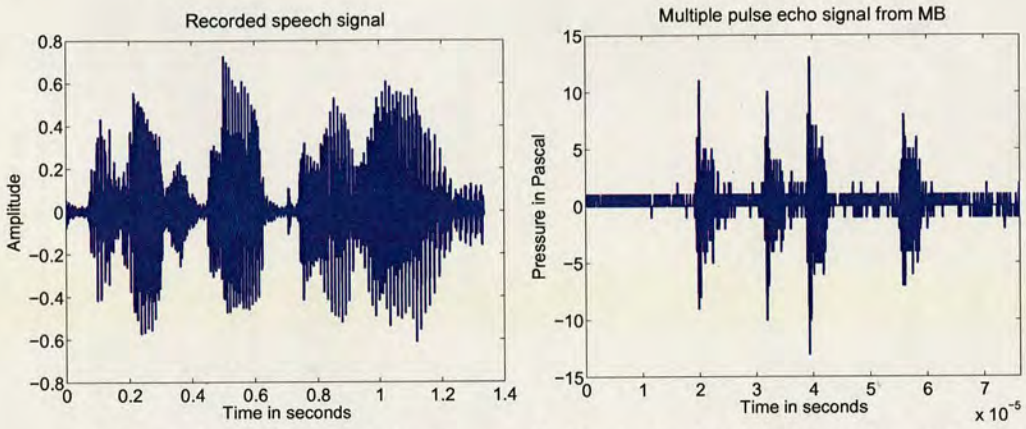
(b) Normalised detected envelope using the Hilbert transform.

Figure 3.10: *Detected envelope of the denoised signal using Hilbert transform.*

pulse locations in the time domain.

From intuitive observation of experimentally measured ultrasound MB echo signals, there are several pulse segments in the signal in the time domain. It is analogous to a typical recorded speech utterance. Speech is a consecutive sequence composed of voiced and unvoiced speech sounds. The voiced sounds are periodic in structure with a fundamental frequency and several harmonics whereas the unvoiced sounds are usually regarded as noise. The voiced sounds originate from the excitation for the vocal tract, which is created by the modulation of the airflow passing through the vocal folds. The vocal track is often modeled as an acoustic tube with resonances called formants, which contain the associated fundamental frequency F_0 and many harmonics F_i [74]. By comparison, since the measured ultrasound incident pulses are composed of six period sinusoids, the corresponding echoes from the scatterers, either SCSs or MBs, are likely to have periodic structures. Both of them can be assumed short-term stationary in the time domain. Moreover, the frequency components, especially the harmonics, are one of the most important features of echo returns in the frequency domain. Figure 3.11(a) displays a typical recorded speech signal and Figure 3.11(b) displays a typical acquired MB echo for comparison. Due to the similarities discussed above, methods for detection and estimation of speech sounds can also be applied to measured ultrasound echo signals.

VAD is commonly used in speech processing [75–77], especially in speech recognition, for tackling the problem of classifying speech and non-speech sounds in a noisy



(a) A typical recorded speech signal.

(b) Received echo signal from the ultrasound MBs.

Figure 3.11: *Comparison of waveform of the speech signal and echo signal from the ultrasound MBs.*

signal. During the past decades, there are enumerate strategies for detecting the speech sound. A complete review of the state of art and the normally used evaluation frameworks are reported in [78]. It has been employed in several applications in the literature including mobile communications [79], real-time speech transmission on the Internet [80] and noise reduction for digital hearing aid devices [81]. Since the VAD algorithm has a good performance in detecting the speech sounds in a signal, it is able to detect the envelope of ultrasound echo returns due to their substantial similarities in both the time and the frequency domains.

Most of the traditional VAD algorithms are mainly based on general speech properties, such as energy thresholds, pitch detection, periodicity measures [78]. They are very sensitive to the SNR and thus may fail when the level of noise increases. Recently developed VAD algorithms are based on a statistical model [77, 82]. In these algorithms, a signal model is adopted, a likelihood ratio is developed and a statistical hypothesis test is conducted. Davis [83] investigated and compared a number of statistical VAD schemes, and also proposed a novel method using an SNR measure. This approach introduced a low-variance spectrum estimate and determined an optimal threshold based on the estimated noise statistics. In the process of this VAD scheme, the original signal, $x(n)$, is first divided into K frames, $x_1(n), \ldots, x_K(n)$. For each frame, $x_k(n)$ $(k = 1, \ldots, K)$, the technique calculates an SNR measure, $\phi_k(f_l)$, and an adaptive threshold, $\eta_{T,k}(f_l)$, for each spectral bin, $f_l \in \{0, 1, \ldots, L-1\}$, and L is the

number of spectral bins. The SNR measure, $\phi_k(f_l)$, is obtained by computing the ratio of the signal power spectral density (PSD) and the noise PSD as:

$$\phi_k(f_l) = \frac{P_{xx,k}(f_l)}{\hat{P}_{vv}(f_l)} - 1, \tag{3.12}$$

where $P_{xx,k}(f_l)$ is the signal PSD of the current frame k and $\hat{P}_{vv}(f_l)$ is the estimated value of the noise PSD, which can be calculated as the sample mean over an initial non-speech period:

$$\hat{P}_{vv}(f_l) = \frac{1}{V}\sum_{k=0}^{V-1} P_{xx,k}(f_l), \tag{3.13}$$

where V is the number of frames during the initial period. Moreover, the adaptive threshold in each spectral bin, $\eta_{T,k}(f_l)$, depends on the variance of the corresponding SNR measure in non-speech periods, $\sigma^2_{v,k}(f_l)$. To be more specific, two hypotheses, H_0 and H_1, are considered in deciding whether the speech is present or not. H_0 represents the case when there is only noise present and H_1 represents the case when there are both speech and noise present. It is also assumed that the SNR measure in non-speech periods is zero mean and Gaussian distributed for hypothesis H_0. Therefore, the probability density function (PDF) of the SNR measure for non-speech activity can be modeled as:

$$p(\phi_k(f_l)|H_0) = \frac{1}{\sqrt{2\pi}}\sigma^2_{v,k}(f_l)\exp\left(-\frac{\phi^2_k(f_l)}{2\sigma^2_{v,k}(f_l)}\right). \tag{3.14}$$

If the adaptive threshold, $\eta_{T,k}(f_l)$, is smaller than the SNR measure, $\phi_k(f_l)$, given the null hypothesis H_0 is accepted, the false alarm is realized. The probability of a false alarm, P_{FA}, is expressed as:

$$Pr(\eta_{T,k}(f_l) < \phi_k(f_l)|H_0) = \int_{\eta_{T,k}(f_l)}^{+\infty} \frac{1}{\sqrt{2\pi}}\sigma^2_{v,k}(f_l)\exp\left(-\frac{\phi^2_k(f_l)}{2\sigma^2_{v,k}(f_l)}\right) d\phi_k(f_l). \tag{3.15}$$

After some simple manipulation of (3.15), as stated in [83], the adaptive threshold $\eta_{T,k}(f_l)$ can be given by:

$$\eta_{T,k}(f_l) = \sqrt{2\sigma^2_{v,k}(f_l)} \cdot \mathrm{erfc}^{-1}(2P_{FA}), \tag{3.16}$$

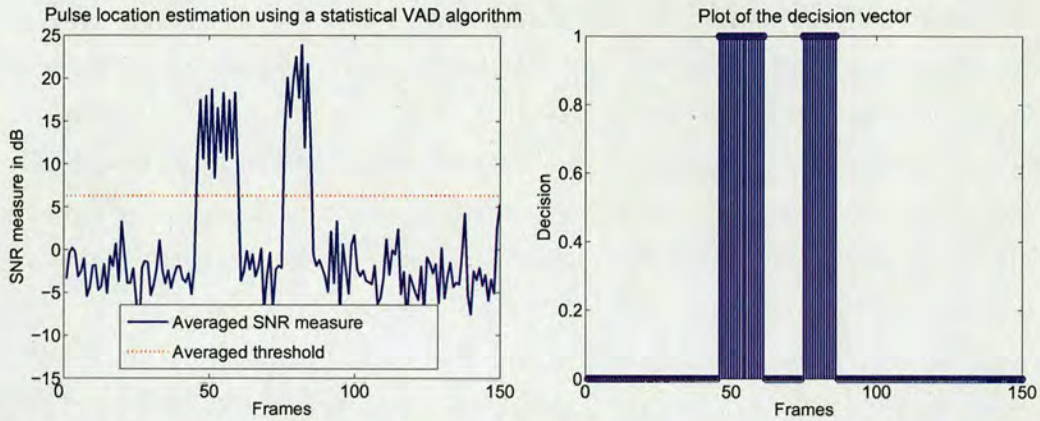
where $\mathrm{erfc}^{-1}(\cdot)$ is the complementary error function. Once $\phi_k(f_l)$ and $\eta_{T,k}(f_l)$ in each

Algorithm 3.1: The procedure of the statistical VAD algorithm

```
Divide the original signal into K frames ;
for k = 1 to K frames do
    for f_l = 1 to L bins do
        Calculate the expected noise PSD P̂_{vv,k}(f_l) ;
        Calculate the signal PSD: P_{xx,k}(f_l) ;
        Calculate the SNR measure of the current frame k: φ_k(f_l) = P_{xx,k}(f_l) / P̂_{vv,k}(f_l) − 1 ;
        if the signal is in the noise only period then
            Calculate the noise variance: σ²_{v,k}(f_l) = [P_{vv,k}(f_l) / P̂_{vv,k}(f_l) − 1]² ;
        end
        Calculate the threshold of the current frame k given a certain probability
        of false alarm P_FA: η_{T,k}(f_l) = √(2σ²_{v,k}(f_l) erfc⁻¹(2P_FA)) ;
    end
    Calculate the averaged threshold η_{T,k} and the averaged SNR measure φ_k
    over all frequency bins for each frame k ;
    Compare η_{T,k} and φ_k to give a decision: 1 if φ_k is larger; 0 if η_{T,k} is larger ;
end
Construct a decision vector D(k) containing 0s and 1s ;
Expand D(k) over the full signal length to determine the start and end points of
each pulse ;
```

spectral bin f_l are calculated, their means are computed over all frequency bins to provide the averaged SNR measure, ϕ_k, and the averaged threshold, $\eta_{T,k}$, respectively. Then they are compared to give a final decision in frame k. If ϕ_k is larger than $\eta_{T,k}$, the signal in the kth frame is regarded as being in the speech period, which is denoted as 1 in the final decision vector $D(k)$. Otherwise, assign 0 to the kth value in the decision vector, which indicates non-speech period is present. Finally, all VAD decisions from K frames in the time domain are in the decision vector $D(k)$. The changes from 0s to 1s and from 1s to 0s in $D(k)$ imply the exact pulse locations. The procedure of this VAD scheme for a signal is illustrated in Algorithm 3.1.

Apply this statistical VAD algorithm to the simulated signal, which is described in Appendix A. The signal in the time domain was shown in Figure 3.6(a) on page 40. It is divided into 150 frames and each frame has a length of 10 data points. The number of initial data points used to calculate the expected noise PSD is heuristically set to 5% of the whole length of the signal. Moreover, the probability of the false alarm in the hypothesis is made at 5%, as suggested in [83]. Figure 3.12(a) shows the result of detection. The averaged SNR measure is denoted in blue solid line and the aver-



(a) Original two pulse signal embedded in the noise.

(b) Detected envelope by VAD technique only.

Figure 3.12: *Envelope detection for noisy signal by VAD algorithm.*

aged threshold is in red dotted line. Compared to the detected envelope shown in Figure 3.9(b) on page 46 using the HTWD method, Figure 3.12(a) displays a smoother curve at the top of the envelope whereas more random noise at the bottom. These uncertainties at the bottom of Figure 3.12(a) will have a large influence on determining the edges of pulses. Furthermore, each element in the decision vector, $D(k)$, represents one frame with ten data points, thus each single 0 or 1 will expand to ten 0s or 1s. According to the heuristic experience from measured ultrasound MB echo signals, two assumptions are made. If the length of a pulse is smaller than 20 data points, then the pulse will be regarded as noise. If the distance between two pulses are smaller than 10 data points, they will be regarded as a single pulse. Figure 3.12(b) shows the decision vector that corresponds to the averaged SNR measure and the averaged threshold displayed in Figure 3.12(a). It only shows the result in 150 frames. However, by expanding one element to ten elements with the same values, the final resulting decision vector in 1500 data points, which is the total length of the original signal, can be obtained easily. Therefore, determining each change-point in the final decision vector corresponds to finding where 0 changes to 1 and where 1 changes to 0. This makes the determination of exact pulse locations in the original signal much easier.

3.2.4 Combining the HTWD method and VAD algorithms

Two approaches have been employed to estimate multiple pulse locations in measured ultrasound echo signals in Sections 3.2.1, Section 3.2.2 and Section 3.2.3. Using

the proposed HTWD method, the denoised envelope using the HT is shown in Figure 3.9(b) on page 46. It can be seen that after the denoising process, the envelope of the multiple pulse signal has ripples only at the top. The ripples at the bottom are almost removed. Alternatively, using the VAD algorithms, the detected envelope is displayed in Figure 3.12(a). It is worth noticing that the top of the envelope is much clearer in Figure 3.12(a) than that in Figure 3.9(b) whereas there are more uncertainties at the bottom of the envelope.

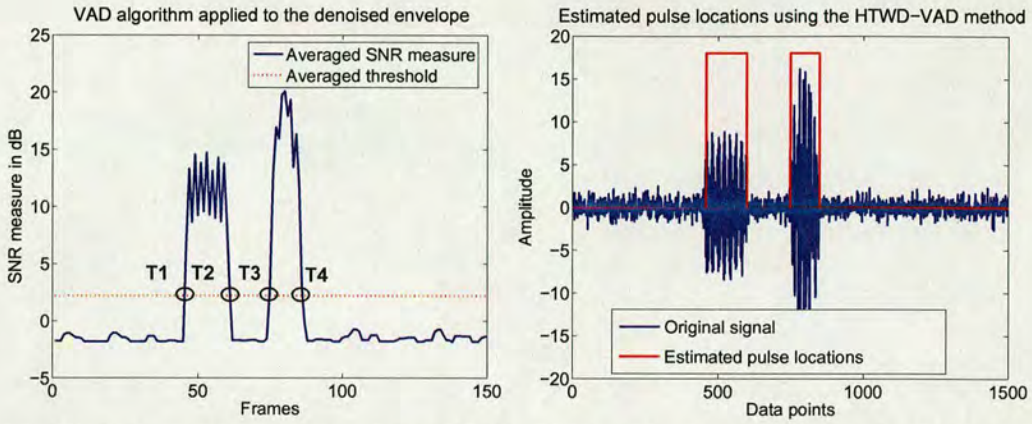
In order to exploit the advantages of both the VAD algorithms and the HTWD method, the combination of two can improve the estimation accuracy, which will be discussed in detail in Section 3.2.4.1. First, the HT is applied to a signal and then the detected envelope is denoised by the WT. Second, a VAD algorithm is performed on the denoised envelope to provide final results for the estimated pulse locations. This newly proposed technique for the estimation of multiple pulse locations is termed as the HTWD-VAD method and the procedure is summarized in Algorithm 3.2.

Algorithm 3.2: The procedure of the HTWD-VAD method for the estimation of multiple pulse locations

1 Envelope detection using the HT ;
2 Denoise the detected envelope using the WT ;
3 Perform a statistical VAD algorithm on the denoised envelope using Algorithm 3.1 ;

3.2.4.1 Evaluation of the HTWD-VAD method on a simulated signal

A simulated signal used for evaluating the newly proposed HTWD-VAD method is presented in Appendix A and has been shown in Figure 3.6(a) on page 40. By taking the advantages of both the HTWD method and the VAD techniques, the HTWD-VAD method produces a smoother curve at both the top and the bottom of the envelope, which is shown in Figure 3.13(a). Compared to Figure 3.9(b) on page 46 and Figure 3.12(a), Figure 3.13(a) provides the clearest envelope with obvious mainlobes and almost flat sidelobes. Thereby it is much easier to determine the start and end points of each pulse with the averaged threshold, which is denoted by a red dotted line in Figure 3.13(a). According to the intersection points of the averaged SNR measure and the averaged threshold, denoted by T_1, T_2, T_3 and T_4 in Figure 3.13(a), Figure 3.13(b) shows the corresponding estimated multiple pulse locations in the time domain.



(a) Detected envelope by the HTWD-VAD method.

(b) The estimation result of the pulse locations.

Figure 3.13: *The estimated results using the proposed HTWD-VAD method.*

This proposed HTWD-VAD method is evaluated over 100 realisations of the simulated signal, which have different noise variances. The mean and the standard deviation (SD) of the estimation results for pulse locations are calculated in data points and compared to the ground truths. Table 3.1 compares means and SD values of the estimated results for pulse locations in the time domain using three approaches: the HTWD method, the VAD method and the HTWD-VAD method. Figure 3.14 illustrates the calculated MSE values in decibel (dB) versus different SNRs using three approaches. The blue dot-marked line, the red circle-marked line and the green diamond-marked line represent the performance of the HTWD method, the VAD method and the HTWD-VAD method respectively. It is shown that when the SNR is lower than 15dB, the HTWD method has better performance than the other two. However, when the SNR is higher than 15dB, the HTWD-VAD method has the best performance. Moreover, the HTWD-VAD method operates the VAD algorithm on the denoised signal envelope and chooses the threshold adaptively according to the noisy signal statistics.

Nevertheless, all three approaches have similar limitations. If two pulses in the echo signal are closely-spaced, these methods cannot distinguish with each other and may regard them as one single pulse. Or if the amplitude of the pulse is relatively small, the methods will consider it as noise. They are due to the assumptions when a threshold is chosen, no matter it is fixed in the HTWD method or adaptive in the VAD method and the HTWD-VAD method.

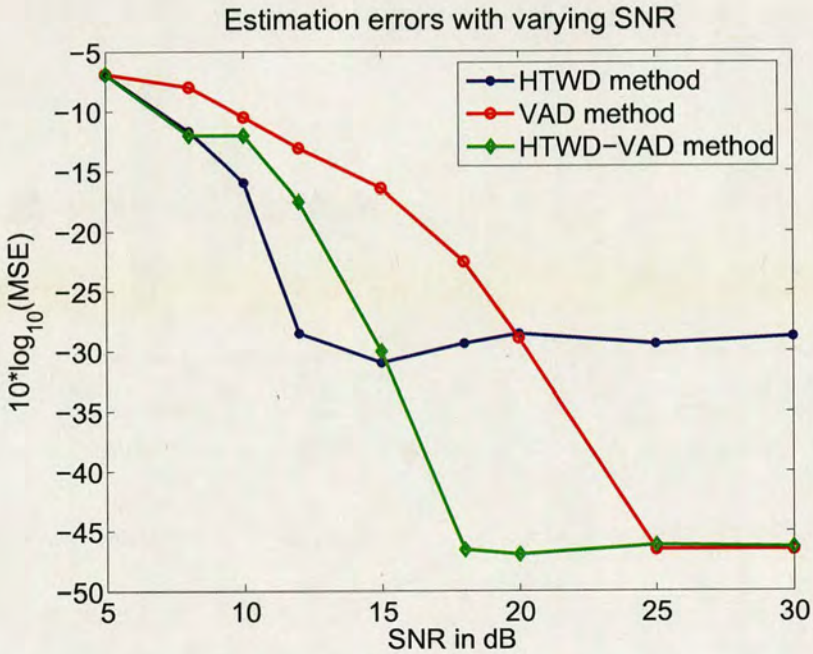


Figure 3.14: *MSE error analysis of the combination method versus SNR.*

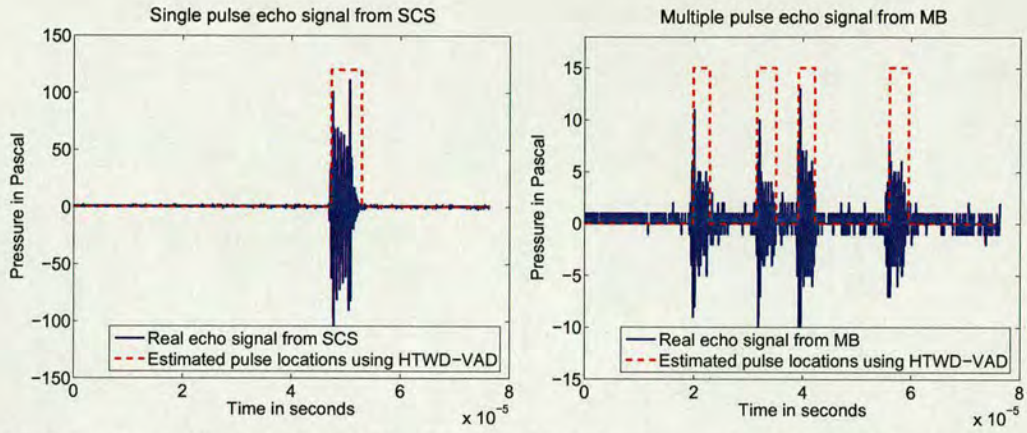
Methods	$\hat{T}_1$ (mean $\pm$ SD)	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_4$
Ground Truth	450	600	750	850
HTWD	395 ± 16	608 ± 1	705 ± 2	936 ± 14
VAD	465 ± 13	571 ± 41	689 ± 105	768 ± 120
HTWD-VAD	460 ± 1	602 ± 4	760 ± 2	850 ± 1

Table 3.1: *Comparison of three estimation approaches for pulse locations in the time domain.*

3.2.4.2 Estimation on measured ultrasound echo signals

After the HTWD-VAD method is evaluated on a simulated signal with ground truths, it can be applied to measured echo signals from ultrasound SCSs and MBs. Take one echo signal from SCSs and one from MBs as examples. They were both acquired from the measured deterministic experiments. The SCS echo and the MB echo were collected with a peak negative pressure of 550kPa and a transmit frequency at 1.83MHz.

The estimated results of pulse locations for the SCS echo signal and the MB echo signal are displayed in Figure 3.15(a) and Figure 3.15(b) respectively. The original echo signals are denoted in blue solid lines and the estimated pulse locations are denoted in red dashed lines. Moreover, the estimated start and end points of each pulse are



(a) Estimation results for a single pulse SCS echo signal.

(b) Estimation results for a multiple pulse MB echo signal.

Figure 3.15: *Estimated pulse locations in the time domain for the measured ultrasound echo signals.*

Pulse segment	1st
Estimates	[930, 1040]

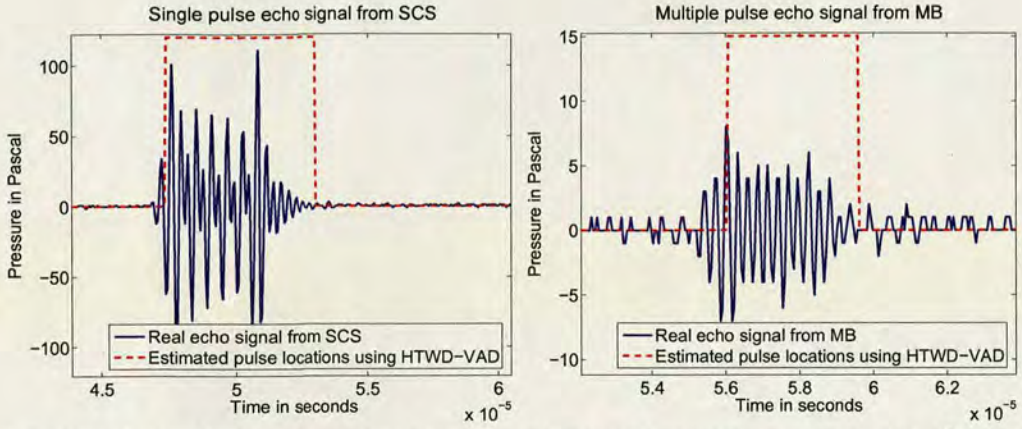
Table 3.2: *Estimated pulse locations of a measured ultrasound SCS echo signal*

Pulse segment	1st	2nd	3rd	4th
Estimates	[390, 450]	[620, 690]	[770, 830]	[1100, 1170]

Table 3.3: *Estimated pulse locations of a measured ultrasound MB echo signal*

displayed in Table 3.2 and Table 3.3 respectively.

Since the results in Figure 3.15 are difficult to see, Figure 3.16 provides an enlarged version for both the SCS echo and the MB echo. In Figure 3.16(a), the estimated results of the single pulse in the SCS echo match well to the observation. In Figure 3.16(b), the enlarged version of the estimates of the 4th pulse segment in the MB echo is shown. Unfortunately, the estimated result of the pulse locations does not match the observation well. It loses some information at the beginning of this pulse. This may result in losing some frequency information or may lead to the misidentification of frequency components in the corresponding pulse segment in the spectral analysis in Section 3.3.



(a) Enlarged estimation results for the single pulse segment in the SCS echo signal.

(b) Enlarged estimation results for the 4^{th} pulse segment in the multiple pulse MB echo signal.

Figure 3.16: *Enlarged estimated pulse locations in the time domain for the measured ultrasound echo signals.*

3.3 Estimation of frequencies in the frequency domain

In Section 3.2, the estimation focused on multiple pulse locations in the time domain. In this section, the estimation concentrates on the spectral analysis in the frequency domain. *Non-parametric* methods for estimation in the frequency domain are generally based on the DFT of either a signal segment or its autocorrelation sequence [51]. If a signal segment is deterministic or it is a wide-sense stationary random process, the PSD exists and thereby can be used to describe how the power is distributed with frequency. Therefore, the power spectrum is able to estimate the spectral content of a stationary signal in the frequency domain. Periodograms and their variations [51], together with the multitaper method [51,84], are all popular choices for computing the power spectra. They will be discussed in Section 3.3.2 and Section 3.3.3 respectively. All these techniques are only valid for one block of signal since both the DFT and the periodograms do not localize in time. If there are multiple pulse segments in a signal, i.e. the signal is time-varying, the STFT can be used as an intuitive method to represent a signal in both the time and the frequency domain. It will be briefly addressed in Section 3.3.4.

3.3.1 Fourier analysis

Fourier analysis is the most commonly used *non-parametric* conventional technique for spectral analysis [51]. The essence of it lies in the representation of a signal as

Window function	Mainlobe width
rectangular	$1.81\pi/(N-1)$
Hanning	$5.01\pi/(N-1)$
Hamming	$6.27\pi/(N-1)$

Table 3.4: *The spectral resolutions with various window functions in the DFT-based Fourier analysis.*

a superposition of sinusoidal components. Fourier analysis is a decomposition of a function into sinusoids of different frequencies. For a deterministic signal, it is able to compute its frequency spectrum or the power spectrum using the Fourier series or Fourier transform.

Since the measured ultrasound echo signals have already been sampled and thus are discrete, there is no need to take the sampling operation and no need to worry about the aliasing error. When a signal is not periodic and not infinite in the time domain, windowing operation is required before applying the process of the DFT.

Using different window functions, there are different degrees of tradeoff between resolution (main lobe width) and the leakage (peak side lobe level). The width of the main lobe is usually specified as the width between the points where the power has fallen -3dB below the maximum value. The spectral resolutions of the DFT using three different commonly used window functions are compared in [51]. Table 3.4 shows this comparison, in which N is the full length of a time domain signal.

After window functions are applied, the DFT can be used to calculate the frequency spectra of signals. Assume a windowed signal $x(n)$, then its DFT and the corresponding IDFT are expressed in (3.17a) and (3.17b).

$$\hat{X}_N(k) = \sum_{n=0}^{N-1} x(n) e^{-j(2\pi/N)kn} \tag{3.17a}$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{X}_N(k) e^{j(2\pi/N)kn}. \tag{3.17b}$$

Unfortunately, there is one fundamental limitation of the Fourier analysis: it cannot localize in time. To be more specific, when analysing a real time non-stationary signal, whose properties are not known in advance, the performance of the Fourier analysis will be very poor. Moreover, the spectral resolution of the DFT only depends on the

length of a signal, which is not satisfactory in many real world applications.

3.3.2 Power spectrum estimation using periodograms

The discussion in Section 3.3.1 focused on the calculation of frequency spectra based on the FT. This section will adopt an alternative approach based on the power spectrum computation to estimate spectral content of a signal in the frequency domain. The periodogram is first introduced by Schuster in 1898 to estimate the power spectrum of the solar sunspot data [51, 85]. The periodogram and its modified versions are basic analysis tools employed in the estimation of power spectrum of stationary signals. The periodogram is defined on a discrete time signal segment $\{x(n)\}_0^{N-1}$. It can be computed using the discrete-time FT (DTFT) of the signal $x(n)$ in (3.18a) directly or using the autocorrelation sequence $\hat{r}_x(l)$ in (3.18b) indirectly [51].

$$\hat{R}_x(e^{j\omega}) \triangleq \frac{1}{N}\left|\sum_{n=0}^{N-1} w(n)x(n)e^{-j\omega n}\right|^2, \tag{3.18a}$$

$$\hat{R}_x(e^{j\omega}) \triangleq \sum_{l=-(N-1)}^{N-1} \hat{r}_x(l)e^{-j\omega l}. \tag{3.18b}$$

$w(n)$ represents a window function in (3.18a). If the window function is rectangular, the estimated power spectrum is termed as *periodogram*; if nonrectangular windows are used, the estimated spectra are termed as *modified periodograms*. Since the measured ultrasound echo signals are already in a discrete set, the values of the periodogram at discrete frequencies can be calculated as:

$$\hat{R}_x(k) \triangleq \hat{R}_x(e^{j2\pi k/N}) = \frac{1}{N}\left|\hat{V}(k)\right|^2 \quad k = 0, 1, \ldots, N-1, \tag{3.19}$$

where $\hat{V}(k)$ is the N-point DFT of the windowed signal segment $v(n) = \omega(n)x(n)$.

It is pointed out in [51] that the periodogram is not an unbiased nor a consistent estimator of the power spectrum of a stationary random signal, therefore increasing the length of signal cannot improve the estimation performance. However, there are two directions for reducing the variance so as to improve the performance of a simple periodogram: *periodogram smoothing* and *periodogram averaging*. These approaches can provide consistent and asymptotically unbiased estimates. The former is based on averaging contiguous values of a single periodogram [86]. The latter is based on splitting a signal into several segments and averaging their respective modified periodograms,

Approach	Spectral resolution
Basic periodogram	$\propto \frac{1}{N}$ (N is the signal length)
Single-periodogram smoothing	$\propto \frac{1}{L_{PS}}$ (L_{PS} is the overlapping length)
Multiple-periodogram averaging	$\propto \frac{1}{L_{PA}}$ (L_{PA} is the segment length)

Table 3.5: *The spectral resolutions for various approaches based on periodograms.*

which can be achieved from multiple nonoverlapping [87] or overlapping of the signal segments [88]. These modified periodograms are able to highly reduce the variance of the estimated power spectrum. However, the spectral resolution will decrease since windowing increases smoothing of peaks. Table 3.5 presents the resolutions of the basic periodogram, the single-periodogram smoothing and the multiple periodogram averaging respectively, as given in [51].

3.3.3 Mutlitaper power spectrum estimation

There is an alternative approach for producing a periodogram-based spectral estimator proposed by Thomson in 1982 [84]. In this method, several data windows are applied to the same data to calculate several corresponding modified periodograms. These modified periodograms are then averaged to produce the estimated power spectrum. The resolution of this spectrum is proportional to $1/N$ [51] and N is the length of the signal. As taping is another name for windowing operation in the time domain, the resulting averaged spectrum is termed as multitaper spectrum.

One significant premise of this approach is that data tapers are required to be properly designed orthogonal functions. Consider a signal segment, $\{x(n)\}_{n=0}^{N-1}$, with a length of N. There is a set of K orthogonal data tapers, $w_k(n)$, where $0 \leq l \leq N-1, 0 \leq k \leq K-1$. They satisfy:

$$\sum_{n=0}^{N-1} w_k(n) w_l(n) = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} . \tag{3.20}$$

Therefore, the averaged multitaper estimator is defined as:

$$\hat{R}_x^{(MT)}(e^{j\omega}) \triangleq \frac{1}{K} \sum_{k=0}^{K-1} \hat{R}_{k,x}(e^{j\omega}), \tag{3.21a}$$

where $\hat{R}_{k,x}(e^{j\omega})$ represents the periodogram estimator that uses the kth taper. By substituting (3.18a) into (3.21a), the estimated multitaper spectrum can be achieved as:

$$\hat{R}_x^{(MT)}(e^{j\omega}) = \frac{1}{KN} \sum_{k=0}^{K-1} \left| \sum_{n=0}^{N-1} w_k(n) x(n) e^{-j\omega n} \right|^2 . \tag{3.21b}$$

Using this approach, the produced periodograms can be considered to be independent at each frequency. Therefore, the averaging would reduce the variance without losing resolution induced by smoothing across frequencies. Moreover, the properly designed full-length windows would reduce bias, which is demonstrated in [51].

3.3.4 STFT for multiple pulse signals

The frequency estimation techniques discussed in previous sections are only valid for stationary signals. Unfortunately, measured ultrasound echo returns are all time varying signals. Approaches, which can be applied to non-stationary signals and can localize in both the time and the frequency domains, are required. STFT is a Fourier-related transform for estimating frequencies in a small fraction of a time-varying signal. It has been mentioned as one of the denoising methods in Section 3.2.2.1. From the perspective of frequency estimation, the detailed description about how the STFT is performed is as follows: a signal is divided into several segments of data. Then the DFT is applied to each single data segment to provide a frequency spectrum in the corresponding time duration. Finally, the individual DFTs from these multiple segments are able to create a 2D plot, which gives an indication of time-frequency properties of the signal. The data segments can be either overlapped or non-overlapped. The overlap of data segments is able to reduce artifacts at the boundary and thereby result in a more accurate time-frequency spectrum. In the case of a discrete time signal, the definition of discrete time STFT is expressed:

$$\mathbf{STFT}\{x(n)\} = X(m,\omega) = \sum_{n=-\infty}^{+\infty} x(n) w(n-m) e^{-j\omega m}, \tag{3.22}$$

where $x(n)$ is the full length signal and $w(n)$ is the window function. The corresponding spectrogram is obtained by computing the magnitude squared of the STFT:

$$\text{spectrogram}\{x(n)\} \triangleq |\, X(m,\omega) \,|^2 . \tag{3.23}$$

In the STFT, both the frequency resolution, i.e. the minimum frequency difference that can be distinguished, and the time resolution, i.e. the time at which frequencies change, are fixed. The width of the window, $w(n)$ in (3.22), i.e. the size of the data segment, is related to the determination of a good time resolution or a good frequency resolution. Specifically, a narrow window renders good time resolution but poor frequency resolution; a wide window provides a good frequency resolution whereas a poor time resolution. Therefore, a trade-off between the time resolution and the frequency resolution needs to be taken into consideration. A number of techniques have been proposed to tackle the problem of fixed resolution. One of the most popular techniques is MRA, which is based on the CWT or the DWT. They can provide good time resolution in high frequency circumstances and good frequency resolution in low frequency circumstances. One application of wavelets in denoising has been briefly discussed in Section 3.2.2.2. Detailed explanations can be found in [66,69]. From many practical perspective, although the wavelets provide a varied resolution for time series signals, they cannot be used as direct replacement of Fourier based techniques due to their intensive computations [89].

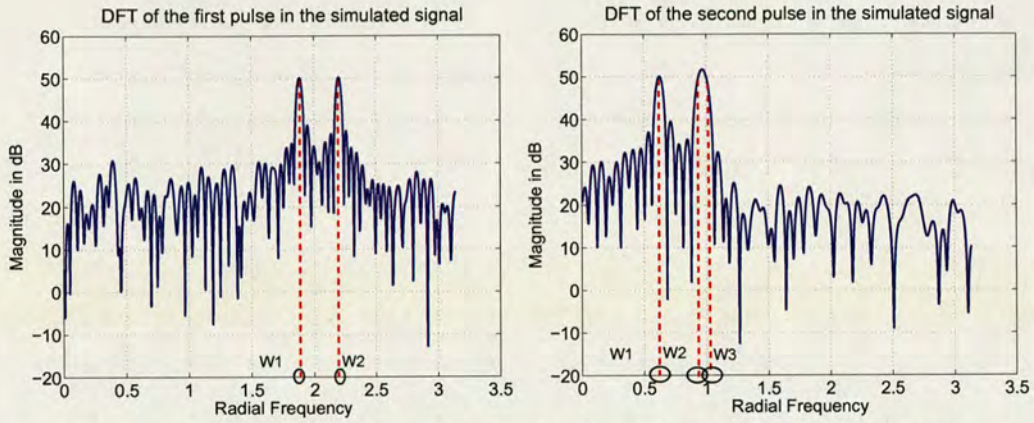
3.3.5 Estimation results for a simulated signal and measured ultrasound echo signals

In this section, three representative frequency estimation techniques are applied to both a simulated signal and measured ultrasound echo signals. They are the basic DFT spectrum estimation, the multitaper power spectrum estimation and the STFT spectrogram estimation. The advantages and disadvantages of the applications of these approaches are also presented.

3.3.5.1 Estimation for a simulated signal

A simulated signal is generated with two pulse segments. The setting of parameters is given in Appendix A and the signal has been shown in Figure 3.6(a) on page 40. In this signal, there are two frequency components in the first pulse and three frequency components in the second pulse. Moreover, two frequencies in the second pulse are closely spaced, which increase the difficulty in discrimination of them. This simulated signal is chosen because it can show the situation when the failure of the discussed approaches appear.

Since the DFT approach and the multitaper approach are only valid for stationary

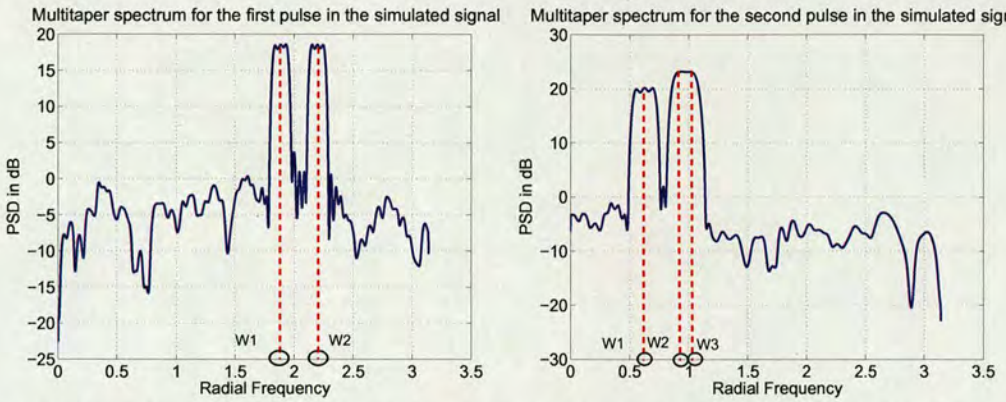


(a) DFT of the 1^{st} pulse in the simulated signal. (b) DFT of the 2^{nd} pulse in the simulated signal.

Figure 3.17: *Magnitude response of frequency spectra using the DFT for the simulated signal.*

signals, the two pulse segments in the simulated signal are separated manually before the estimation techniques are performed. Therefore, the frequency spectra computed by the DFT and the multitaper power spectra can be performed on each pulse segment individually. Figure 3.17(a) and Figure 3.17(b) display the frequency spectra calculated by the DFT for two pulse segments respectively. For the first pulse, two peaks in the frequency spectrum are shown, indicating the detection of two frequency components. However, for the second pulse, in some occasions which depends on the phase of the components, the two closely-spaced frequencies are not resolvable. For example, as shown in Figure 3.17(b), there are two peaks shown in the frequency spectrum of the second pulse and a third peak is missing. Furthermore, the first sidelobe magnitude is about 11dB below the peak of the mainlobe for the first pulse. For the second pulse, the magnitude difference between the lowest peak of the mainlobe and the first sidelobe is about 10dB.

Figure 3.18(a) and Figure 3.18(b) show the multitaper power spectra for the first and the second pulse segments respectively. The two closely-spaced frequencies in the second pulse still cannot be discriminated. There are also only two peaks shown in the multitaper spectrum for the second pulse. ω_2 and ω_3 merge into one peak, as illustrated in Figure 3.18(b). Moreover, the difference between the lowest mainlobe peak and the highest sidelobe peak are about 18dB for both the first pulse and the second pulse, which are much improved compared to Figure 3.17(a) and Figure 3.17(b). However, the widths of the mainlobes in multitaper spectra are wider than those in the frequency



(a) Multitaper power spectrum of the 1^{st} pulse in the simulated signal.

(b) Multitaper power spectrum of the 2^{nd} pulse in the simulated signal.

Figure 3.18: *Multitaper power spectra for the simulated signal.*

spectra.

The techniques discussed above require the simulated signal to be separated into two individual pulse segments before the estimation. By contrast, the STFT approach is able to take the full length of the simulated signal into account as it can localize in both the time and the frequency domains. Figure 3.19 displays the STFT of the full length simulated signal. Hanning window is chosen, of which the length is set to 10 data points and the overlapped segment is set to 90 data points. In the spectrogram, X axis represents the time domain information, which is related to the pulse locations in the signal. Y axis represents the spectral content of the signal. Moreover, the colorbar is given, which indicates the energy scale in the spectrogram. It can be seen that in terms of the time, the energy of the simulated signal are accumulated during two time slots, indicating the time-varying property of the signal. In terms of the radial frequency, in the first time duration, there are two obvious frequency components; in the second time duration, there are also two frequency components present. The third one still cannot be detected.

Although these *non-parametric* techniques for frequency estimation are easy to implement in many applications, none of them can differentiate two closely-spaced frequency components. Furthermore, no matter in the DFT frequency spectrum, or in the multitaper power spectrum, or in the STFT spectrogram, other methods need to be incorporated to determine the exact peak locations. Especially in multitaper power spectrum, the widths of the peaks are wide and in the STFT spectrogram, the fre-

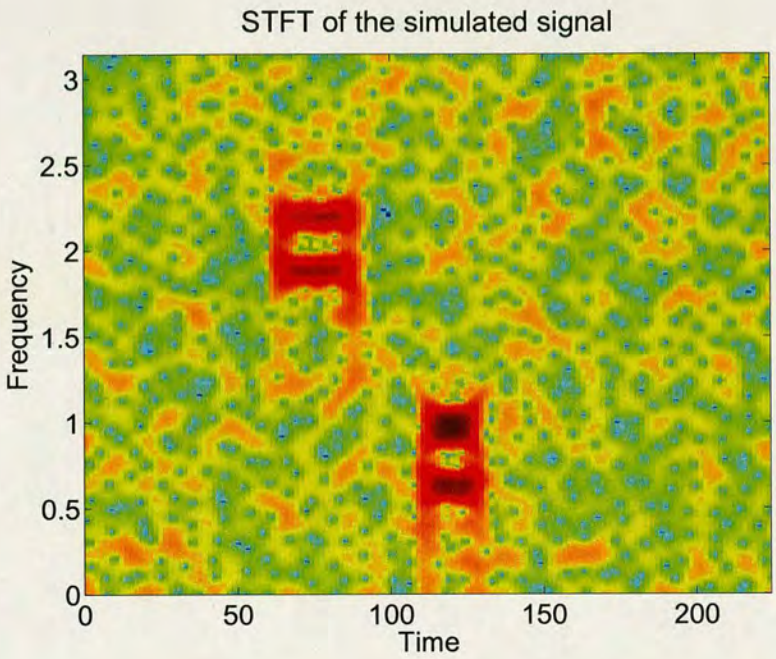


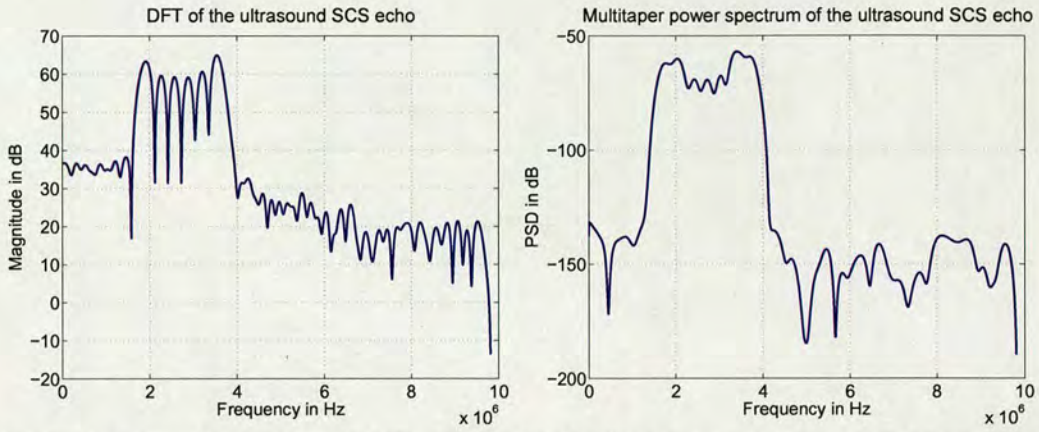
Figure 3.19: *The STFT of the full length simulated signal.*

quency resolution is also poor, which increase the difficulty in determining the peak locations. As a consequence, *non-parametric* spectral estimation approaches only provide a rough understanding about the characteristics of a signal in the time domain and the frequency domain. More advanced techniques with higher accuracy for spectral estimation are required.

3.3.5.2 Estimation for experimentally measured ultrasound echo signals

Non-parametric frequency estimation techniques are applied to the measured ultrasound echo signals in this section to explore their spectral characteristics. The single pulse SCS echo signal was shown in Figure 3.15(a) on page 55, and the multiple pulse MB echo signal was shown in Figure 3.15(b).

Figure 3.20(a) and Figure 3.20(b) illustrate the frequency spectrum and the multitaper power spectrum of the single pulse SCS echo signal respectively. The DFT frequency spectrum has obvious six peaks whereas the peaks in the multitaper power spectrum are wider and are smoothed, thus cannot be easily picked out. All these peaks locate between 2.0MHz and 4.0MHz. The first peak in the DFT spectrum of the SCS response indicates a fundamental frequency at 1.83MHz. Moreover, the peaks in the spectrum seems to be uniformly distributed. Furthermore, the difference between



(a) DFT spectrum of the single pulse SCS echo signal.

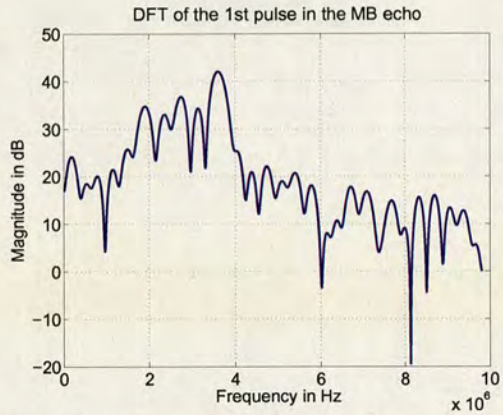
(b) Multitaper power spectrum of the single pulse SCS echo signal.

Figure 3.20: *DFT spectrum and multitaper power spectrum of the measured single pulse SCS echo signal.*

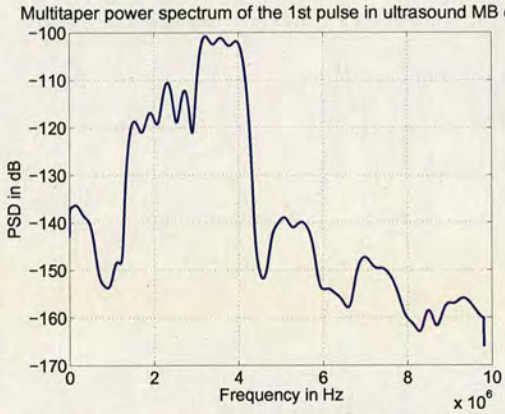
the mainlobe peaks and the first sidelobe in the DFT frequency spectrum is 25dB and the difference between the main lobe peaks and the first side lobe in the multitaper power spectrum is 70dB.

In contrary to the SCS echo signal, there are four pulse segments in the MB echo signal. They are first manually separated, the DFT spectrum and the multitaper power spectrum are then applied to each single pulse segment individually. The estimated results are shown in Figure 3.21 and Figure 3.22. All multitaper power spectra have more flat sidelobes than the DFT spectra. Unfortunately, the mainlobe peaks in the multitaper power spectra are much wider and smoothed as a compromise. Furthermore, the peaks in the DFT spectra or multitaper spectra of the MB response are around the second harmonic frequencies. All four pulse segments in the MB response have similar frequency distributions. They have a tendency that the dominant frequencies are more and more close to the second harmonic, 3.66MHz, from the first pulse segment to the fourth pulse segment. Moreover, less and less frequencies around the fundamental are shown in both the DFT spectra and the multitaper power spectra.

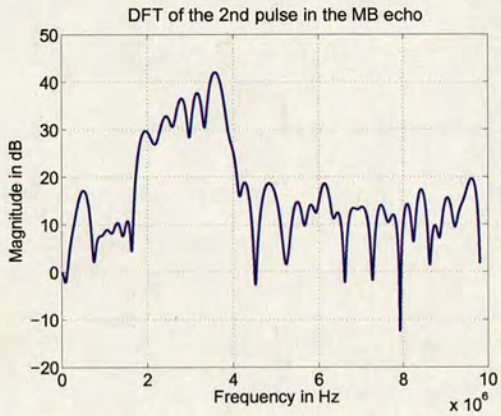
In addition to the DFT spectra and the multitaper power spectra, the STFT spectrograms are also applied to detect both the time information and the frequency information of the measured echo signals. Figure 3.23(a) and Figure 3.23(b) display the STFT spectrograms for the SCS echo signal and the MB echo signal. The window function is Hanning and the length of the window is 200 data points, and the overlapped segment



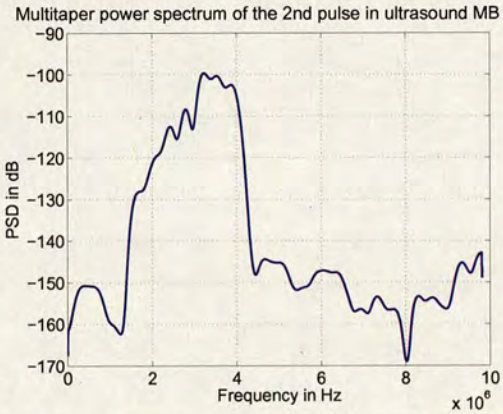
(a) DFT spectrum of the 1^{st} pulse in the MB echo signal.



(b) Multitaper power spectrum of the 1^{st} pulse in the MB echo signal.



(c) DFT spectrum of the 2^{nd} pulse in the MB echo signal.



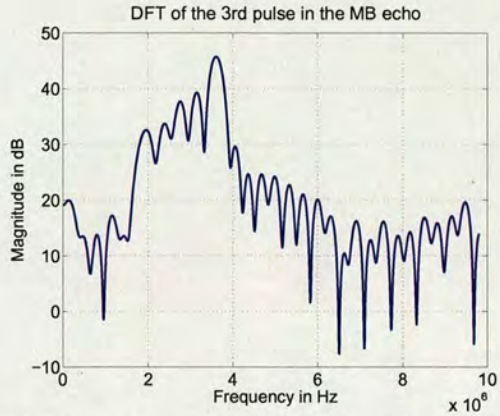
(d) Multitaper power spectrum of the 2^{nd} pulse in the MB echo signal.

Figure 3.21: *Spectral estimation of the 1^{st} and the 2^{nd} pulses in the measured MB echo signal.*

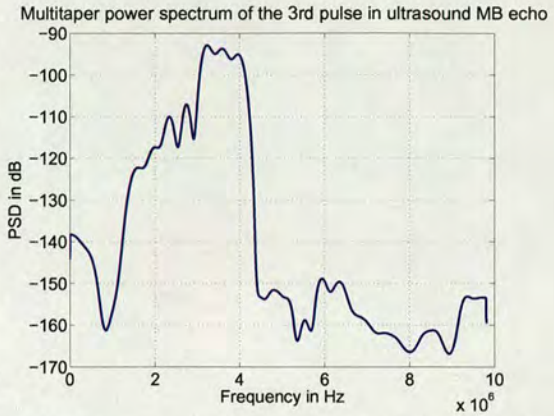
is 190 data points. According to the shown spectrograms, the time domain information and the frequency domain information can be detected simultaneously. However, the resolutions are unsatisfactory, which will make it very hard to determine exact values of the pulse locations and the spectral contents.

3.4 Chapter summary

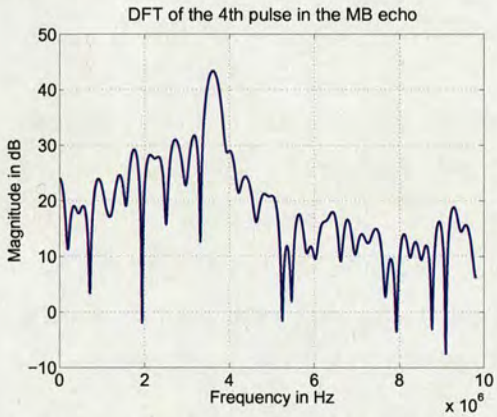
This chapter has introduced several popular techniques currently available in the signal processing field into the analysis of experimentally measured ultrasound echo signals and revised them to be suitable for use in the ultrasonic scenario. These preliminary estimation methods are *non-parametric* in both the time and the frequency



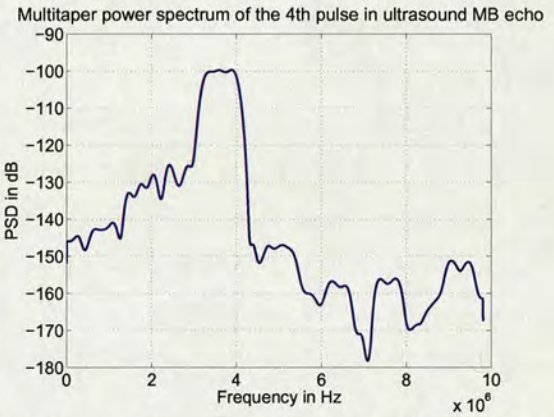
(a) DFT spectrum of the 3^{rd} pulse in the MB echo signal.



(b) Multitaper power spectrum of the 3^{rd} pulse in the MB echo signal.



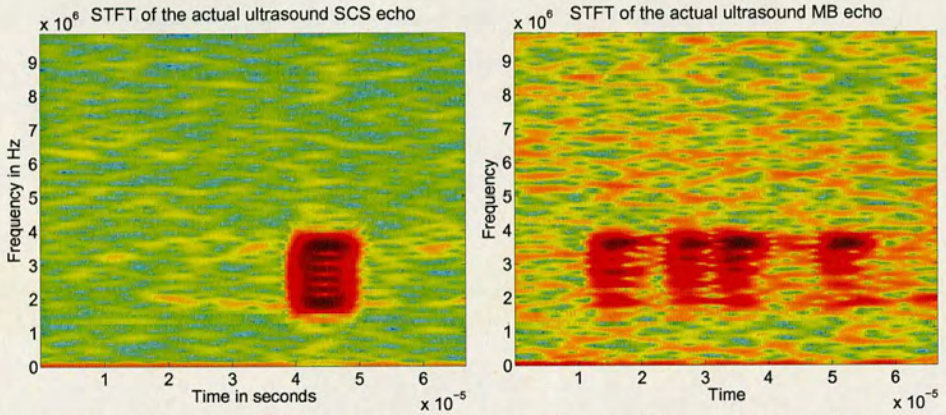
(c) DFT spectrum of the 4^{th} pulse in the MB echo signal.



(d) Multitaper power spectrum of the 4^{th} pulse in the MB echo signal.

Figure 3.22: *Spectral estimation of the 3^{rd} and 4^{th} pulses in the measured MB echo signal.*

domains. In order to estimate the multiple pulse locations of echo signals in the time domain, a novel method is proposed. It combines an envelope detection method and a VAD algorithm. This approach has better performance than that of each single technique. In the frequency domain, the multitaper spectrum estimation has been adopted to estimate the spectral content of echo signals. The side lobes have been suppressed significantly although the resolution of spectrum has degraded compared to the corresponding Fourier spectrum. Apart from these techniques that are only valid for stationary signals, the adopted STFT technique can localize in both the time domain and the frequency domains, although the resolution of them are not satisfactory. Therefore, a more sophisticated *parametric* estimation approach is required so as to improve the estimation performance.



(a) STFT spectrogram of the single pulse SCS echo signal.

(b) STFT spectrogram of the multiple pulse MB echo signal.

Figure 3.23: *STFT spectrograms of the SCS and the MB echo signals.*

Furthermore, *non-parametric* estimation methods in the time domain and in the frequency domain are two independent procedures. If a signal model is employed in the estimation, the process of estimation using a *parametric* method is able to combine the two individual domains and jointly extracts the characteristics of echo returns in the time and the frequency domains. The detailed *parametric* estimation method for measured ultrasound echo signals will be described in Chapter 4 and Chapter 6.

Chapter 4
Parametric modeling and Bayesian inference for measured ultrasound echo signals

Techniques were developed in Chapter 3 for the analysis of measured ultrasound echo signals using *non-parametric* methods while this chapter focuses on *parametric* methods based on signal modeling. Several signal models are first compared according to the physical understanding of measured ultrasound echo signals, which explains why a sum of sinusoids model is chosen. Then a more complicated signal model, which contains both temporal and spectral information of a signal, is proposed. In the process of the *parametric* estimation, Bayesian inference is introduced to estimate model parameters. The application of the inference to the proposed signal model for measured ultrasound echoes is also presented.

4.1 Introduction

Non-parametric methods, for estimating both temporal and spectral contents of measured ultrasound echo signals, have been investigated in Chapter 3. In the time domain, the *non-parametric* estimation methods discussed in Section 3.2 were able to provide good estimates. Unfortunately, they were performed independently and thus were hard to extend to include estimation of other information. From the perspective of spectral estimation, using the classical methods described in Section 3.3, the spectra are derived from the observation data, which is either explicitly or implicitly windowed. It is assumed that the data outside the window is zero. In many practical applications, this is not a reasonable assumption and the window effect will limit the resolution obtained by the estimator [63]. Moreover, most of *non-parametric* spectral estimation methods are only valid for stationary signals.

In contrast, in *parametric* estimation methods, a signal model is assumed to generate all data samples rather than just the data that is actually observed. In this case, no

windowing occurs and the resolution limitations may be overcome. However, the cost of this spectral resolution gain is that the model should fit the data. Therefore, a *parametric* estimation method is demanding when the estimation of time-varying signals with high accuracy is necessary, or when the estimation in both the time and the frequency domains simultaneously is required.

In this chapter, a general discussion about signal modeling and parametric estimation is first addressed in Section 4.2. For the selection of signal model structure, in the particular application of spectral analysis, a sum of sinusoidal model is proposed in Section 4.3.2 and then some other variations of this model are also examined in Section 4.3.3. Compared to measured ultrasound echo signals, an appropriate model for a single pulse segment is proposed in Section 4.3.4.1 and a model which suits for the full length signal is proposed in Section 4.3.4.2. After a signal model has been determined, Bayesian inference is employed to estimate model parameters in Section 4.4. Section 4.5 only applies the inference to the proposed model for measured ultrasound echo signals and gives the likelihood functions for the proposed model. As the posterior distributions for model parameters are difficult to obtain and to sample from, a more detailed investigation of parameter estimation for this particular signal model with some advanced sampling techniques will be examined in Chapter 6. Moreover, the model order selection is another important issue in signal modeling, which is briefly discussed in Section 4.6. Furthermore, the employment of estimators can be used to obtain the estimates for model parameters from a number of data samples. A simple introduction is described in Section 4.7.

4.2 Signal modeling and parametric estimation

The term *model* is usually used to explain or describe the hidden laws that are able to constrain the generation of physical data of interest [90]. For many practical modeling applications, the process of signal model building can be described in the following steps:

1. Choose the model structure.
2. Choose the model order.
3. Estimate the corresponding model parameters.
4. Evaluate the performance of the candidate model.

5. Use the model for real applications.

The first step of model structure selection will be discussed in 4.3. It depends on the physical nature inside a signal. With a known model order, the corresponding model parameters can be estimated. The estimation and evaluation steps should be considered within a specified signal model before being applied to measured signals. Therefore, Section 4.3.4 will propose a specific signal model for measured ultrasound echo signals and Section 4.4 will employ Bayesian inference into model parameter estimation, which are addressed within a fixed model order. The last two steps of evaluation the performance of a candidate model and the use for real applications will be presented in Chapter 6. Moreover, model order selection is another critical issue in *parametric* estimation, which will be discussed in Section 4.6 in detail. As a result, this procedure of five steps provides basic concepts of signal modeling and *parametric* estimation.

4.3 Model structure selection

When choosing a model structure, the understanding of a signal and the physical mechanism that generates the signal are of the most significant importance [51]. For example, when the goal of signal analysis is to reveal frequency information hidden in the data, there are two general signal models: *pole-zero* models and *harmonic* models.

If there is a lack of prior information or a lack of sufficient knowledge of the physical mechanism generating the signals, a preliminary data analysis provides sufficient information to choose a pole-zero (PZ) model and some initial estimates for the parameters to start the model building process. A PZ model, also known as autoregressive moving average (ARMA) random signal model, assumes a linear time-invariant system that is excited by white noise. When estimating spectral content, the parameters of PZ model are estimated using the modified Yule-Walker equations [63]. The PSD of the output signal from the assumed system is then computed based on the estimated parameters. More details about PSD estimation using PZ models are described in [91].

The other choice of the model structure is a *sinusoidal* or a *harmonic* model. Based on these models, an observed data sequence can be represented as complex exponentials contaminated in white noise. There are many algorithms available for spectral estimation of the harmonic models. These techniques are able to resolve spectral content closely-spaced in frequency spectra [51]. Most of them, such as Pisarenko harmonic

decomposition (PHD) [92], multiple signal classification (MUSIC) [93] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [94, 95], are based on the eigen-decomposition of a correlation matrix of the data, and its partitioning into signal and noise subspaces. The eigenvectors associated with the smallest eigenvalues is then used to estimate frequencies of the complex exponentials.

Jaynes [96] introduced the principle of Bayesian inference into spectral analysis for a *sinusoidal* model, which will be discussed in detail in Section 4.4. It explores new insights in the spectral analysis. The frequency resolution of this technique depends directly on the SNR, which usually has demonstrated to have better performance than those of conventional PSD estimation techniques [31]. Moreover, Bretthotst extended Jaynes's work to more complex signal models with additive Gaussian noise in [32].

4.3.1 General signal model

The selection of a model structure highly depends on how much understanding can be obtained from the physical mechanism that generates the signal. In real world measurements, a discrete data set $\mathbf{y} = [y_1, \ldots, y_N]^T$ is recorded. They are sampled at discrete times $[t_1, \ldots, t_N]^T$. Assume a signal model, $y_i = s(t_i) + n_i$ at each time $i = 1, \ldots, N$, where $s(t_i)$ is the signal function that needs to be determined and n_i is the noise. It can also be written in a vector form as:

$$\mathbf{y} = \mathbf{s} + \mathbf{n}, \tag{4.1}$$

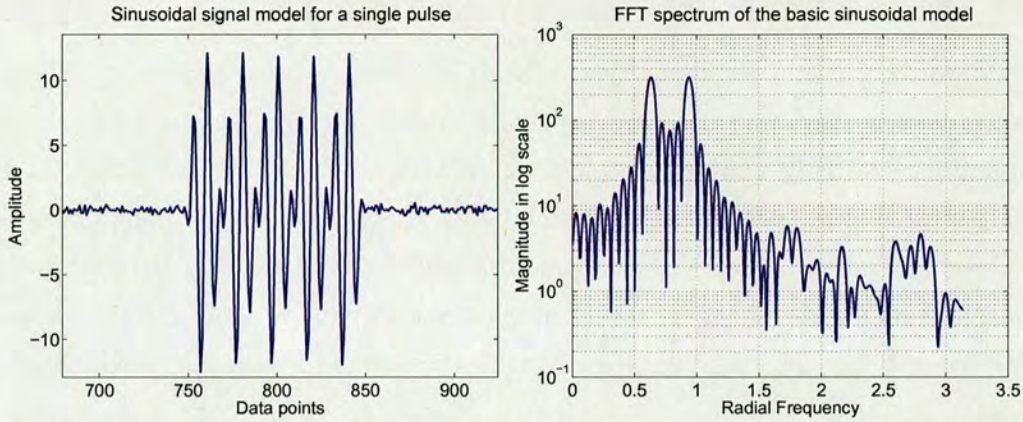
where $\mathbf{s}$ can be generally expressed in a flexible form:

$$\mathbf{s} = \sum_{k=1}^{K} \mathbf{A}_k \mathbf{R}_k(\mathbf{\Psi}). \tag{4.2}$$

In (4.2), the basis function $\mathbf{R}_k(\mathbf{\Psi})$ has a set of parameters $\mathbf{\Psi}$, which can be frequencies, phase shifts and other parameters of interest. $\mathbf{A}_k$ denotes the amplitude vector, which corresponds to the kth basis function $\mathbf{R}_k(\mathbf{\Psi})$.

4.3.2 Sum of sinusoidal model

In various applications, finding a proper basis function and estimating the associated parameters are necessary so as to ensure the model will be a good fit for measured data. Considering the physical mechanism that generates ultrasound echo signals



(a) The basic sinusoidal model in the time domain.

(b) The frequency spectrum of the basic sinusoidal model.

Figure 4.1: *The basic sinusoidal model in the time and the frequency domains.*

k	$E_k \triangleq a_{s,k}^2 + a_{c,k}^2$	$-\arg\tan(a_{s,k}/a_{c,k})$	ω_k
1	40	$\pi/3$	0.2π
2	40	$\pi/4$	0.3π

Table 4.1: *Parameters setup for a single pulse segment in the constructed synthetic signal*

addressed in [2, 3], the basis function in (4.2) can be chosen as Fourier basis, which consists of sines and cosines. A basic sum of sinusoidal model, which contains K frequency components, can be defined in a mathematical way according to (4.1) and (4.2):

$$y(t) = \sum_{k=1}^{K} \{a_{c,k}\cos(\omega_k t) + a_{s,k}\sin(\omega_k t)\} + n(t), \tag{4.3}$$

where ω_k is the kth frequency component in the signal, and $a_{c,k}$ and $a_{s,k}$ are the corresponding amplitudes for the kth frequency.

Figure 4.1(a) displays a simulated single pulse signal constructed by a sum of sinusoidal model embedded in a Gaussian noise. The full length is 1500 data points. The pulse segment locates between $(750, 850)$, which is modeled using (4.3) with $K = 2$. The parameters setup is shown in Table 4.1. The sampling frequency in this simulated signal is 2π. Figure 4.1(b) illustrates its corresponding frequency spectrum using the DFT. It is calculated using the single pulse segment in the signal between the data points 750 and 850, rather than using the full length of signal with 1500 data points.

4.3.3 Other variations of the sinusoidal model

In addition to the sum of sinusoidal model, there are many other popular candidate models that are based on the basic sinusoidal model in (4.3). They are created by multiplying the basic sum of sinusoidal model with different window functions to obtain an amplitude modulated signal model. There are two representative models amongst others: the exponential sinusoidal model [97] and the Gaussian-apodized sinusoidal model [98]. However, these modified models will introduce extra unknown parameters, which also need to be estimated. This will add more complexity to the signal model.

In the case of exponential sinusoidal model (ESM), a signal is modeled using a superposition of time-varying exponentially weighted sinusoids and a general expression is as follows:

$$y(t) = e^{-d(t)} \cdot \sum_{k=1}^{K} \{a_{c,k} \cos(\omega_k t) + a_{s,k} \sin(\omega_k t)\} + n(t), \tag{4.4}$$

where $e^{-d(t)}$ represents an exponentially decaying function.

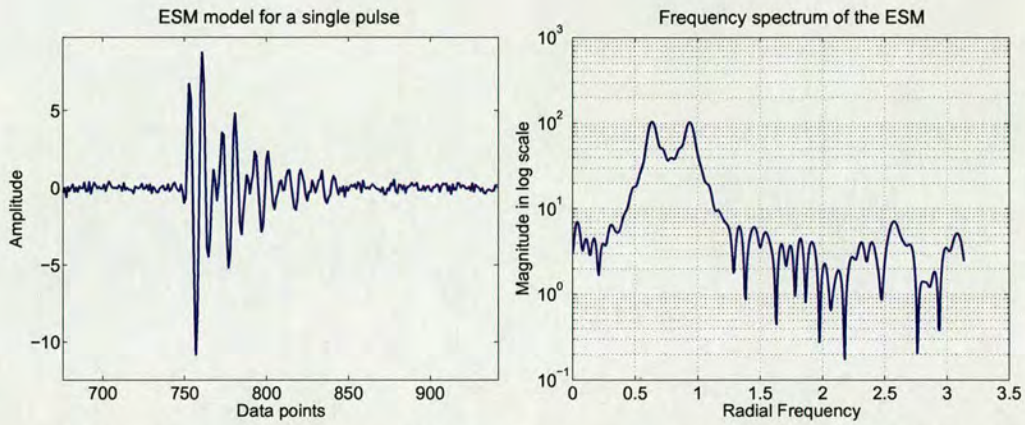
Apart from the ESM, another choice of an amplitude modulated signal model employs a window function to the basic sinusoidal model. There are various window functions available, such as rectangular window, Hanning window and Gaussian window. Since these window functions have similar properties in both the time domain and the frequency domain [51], Gaussian window is chosen as a more flexible and adjustable function. The mathematical expression of a Gaussian windowed signal, which is termed as Gaussian-apodized sinusoidal model (GASM), is given by:

$$y(t) = w(t, \alpha) \cdot \sum_{k=1}^{K} \{a_{c,k} \cos(\omega_k t) + a_{s,k} \sin(\omega_k t)\} + n(t). \tag{4.5}$$

The Gaussian window $w(t, \alpha)$ has a general form of:

$$w(t, \alpha) = \exp\left\{-\frac{1}{2}\left[\frac{t - (N-1)/2}{\alpha(N-1)/2}\right]^2\right\}, \quad \alpha \leq 0.5, \tag{4.6}$$

where α is the standard deviation of the Gaussian window, and N is the window length. Moreover, it is assumed that the Gaussian window is centered on the midpoint of the pulse segment. In the example shown in Figure 4.1(a), the Gaussian window is



(a) ESM for a single pulse in the time domain.

(b) Frequency spectrum of the ESM in the frequency domain.

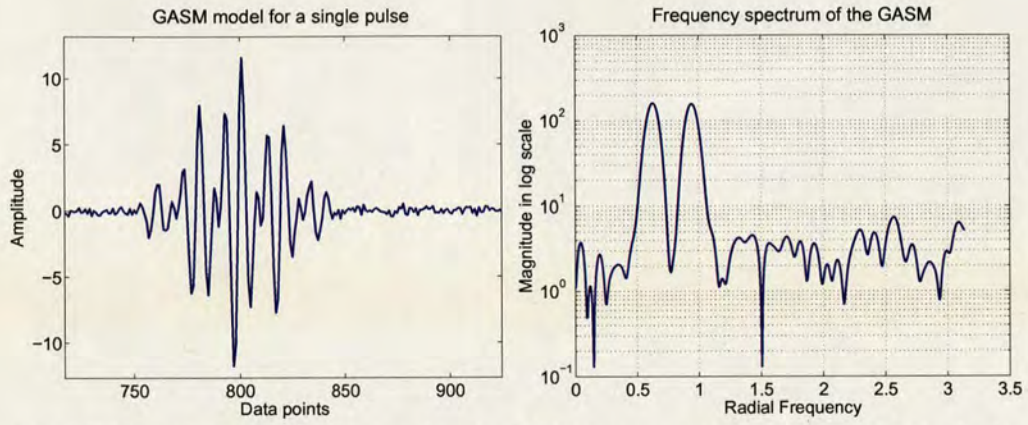
Figure 4.2: *Display the ESM in the time and the frequency domains.*

centered on the data points at 800.

Based on the simulated signal shown in Figure 4.1(a), examples of the ESM and the GASM in the time domain, are constructed. In the ESM, the decaying function is selected as a linear function $d(t) = \lambda_0 t$, in which $\lambda_0 = 3$ is an example of a decaying factor. In the GASM, the Gaussian window also has an extra parameter α, which has an example value of $1/2.5$. According to these amplitude modulation, the modified models of ESM and GASM in the time domain are shown in Figure 4.2(a) and Figure 4.3(a) respectively. Their corresponding frequency spectra are also displayed in Figure 4.2(b) and Figure 4.3(b). Due to the exponential decaying and the Gaussian window, the spectra of both models exhibit variations compared to the basic sum of sinusoidal model.

4.3.4 Proposed signal model for measured ultrasound echo signals

In Section 4.3.2 and Section 4.3.3, a sum of sinusoidal model and its variations have been presented theoretically and then compared on a simulated signal in the case of spectral analysis. This section will investigate all the experimentally measured ultrasound echo signals and propose a most appropriate signal model, which not only takes account the spectral content, but incorporates the temporal information in the echo signals as well.

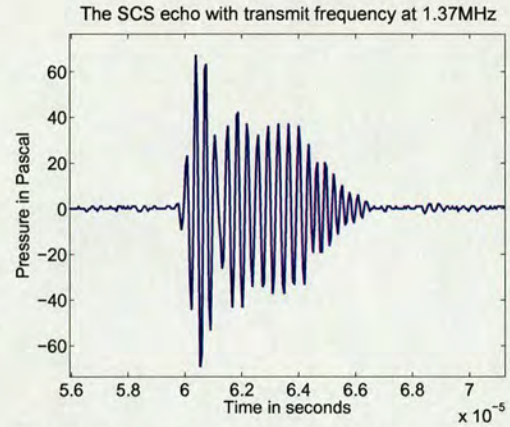


(a) GASM for a single pulse in the time domain. (b) Frequency spectrum of the GASM in the frequency domain.

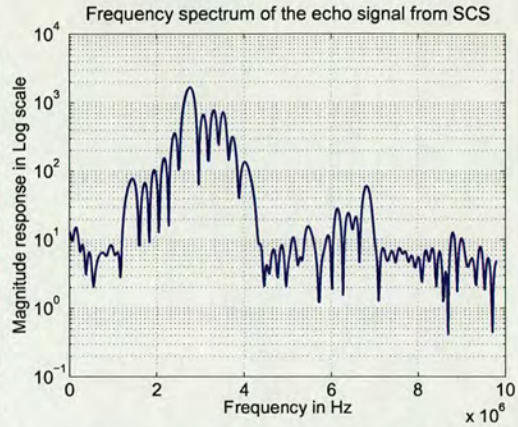
Figure 4.3: *Display the GASM in the time and the frequency domains.*

4.3.4.1 Model structure selection for a pulse segment

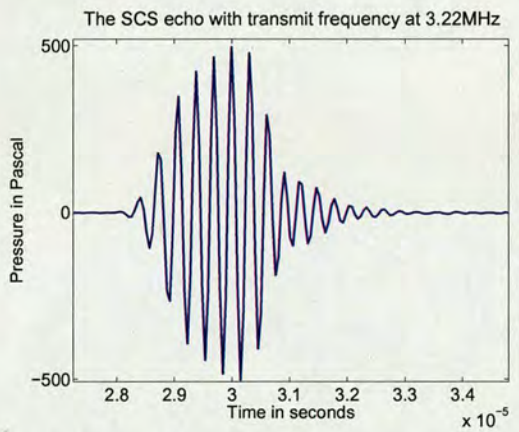
From observation of measured ultrasound echoes from SCSs in the time domain, each pulse segment of the echo signals usually have different shapes with increasing transmit frequency, from 1.2MHz to 4.0MHz. For example, the measured time domain echo signals from SCSs, with the transmit frequency at 1.37MHz and 3.22MHz respectively, are shown in Figure 4.4(a) and Figure 4.4(c). Their corresponding frequency spectra are shown in Figure 4.4(b) and Figure 4.4(d). Compare them with three different models shown in Figure 4.1, Figure 4.2 and Figure 4.3, which have been investigated in Section 4.3.2 and Section 4.3.3. The results of the comparison indicate that, if the transmit frequency is around the lower limit of the frequency range, such as 1.37MHz, the shape of the measured SCS echo is more like an exponentially decaying sinusoidal signal. If the transmit frequency is close to the upper limit of the frequency range, such as 3.22MHz, the measured SCS echo in the time domain is more likely to be regarded as a Gaussian-apodized sinusoidal signal. Nevertheless, the SCS responses with the transmit frequencies between 1.37MHz and 3.22MHz are more like the basic sum of sinusoidal model in the time domain. Take one SCS response with the transmit frequency at 1.83MHz as an example. The time domain echo signal and its corresponding frequency spectrum are shown in Figure 4.5. It is a typical response and thus is representative amongst all measured echo signals from SCSs. Furthermore, the measured ultrasound echo signals are very likely to have many closely-spaced frequency components by intuitive observation in the frequency domain, no matter what the transmit



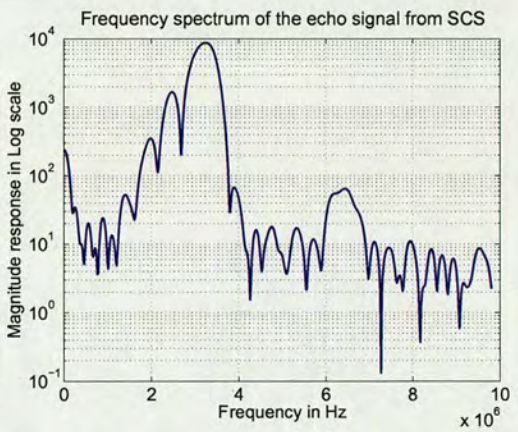
(a) Enlarged SCS response with transmit frequency at 1.37MHz in the time domain.



(b) Frequency spectrum of the SCS response with transmit frequency at 1.37MHz in the frequency domain.



(c) Enlarged SCS response with transmit frequency at 3.22MHz in the time domain.

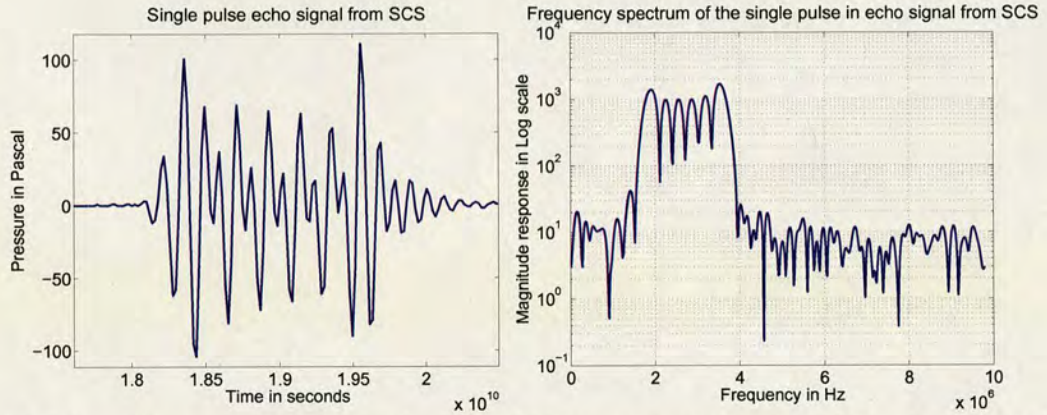


(d) Frequency spectrum of the SCS response with transmit frequency at 3.22MHz in the frequency domain.

Figure 4.4: *The SCS responses with transmit frequency at* 1.37*MHz and* 3.22*MHz.*

frequency is.

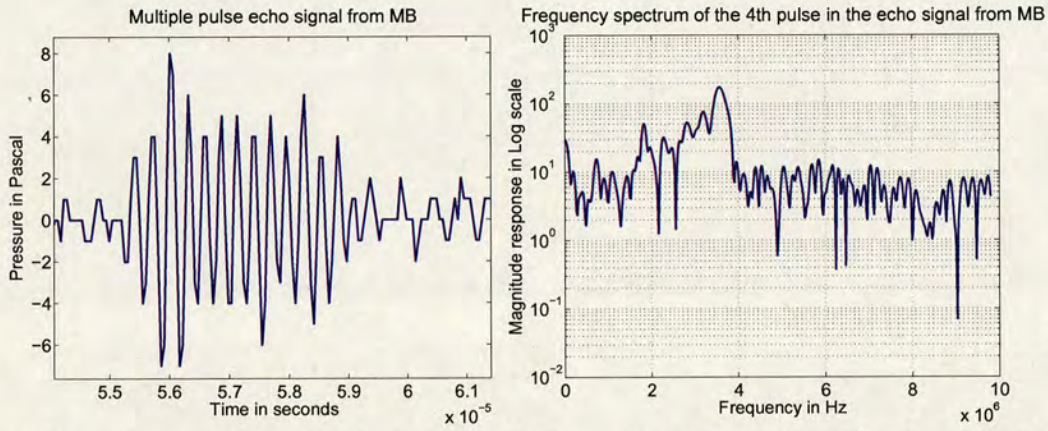
In contrast to SCS responses, most of MB responses have multiple pulses and almost all of them have no amplitude modulation on each single pulse of the signal, no matter what transmit frequency at the center of the transducer. Moreover, the SNR of an MB response is much worse than that of an SCS response. One example of MB responses discussed in the following has the same transmit frequency at 1.83MHz as the SCS response shown in Figure 4.5(a). Figure 4.6(a) illustrates the 4th pulse segment of the chosen MB echo signal. It is shown that the amplitude of the MB response in the time domain is small in comparison to the SCS response. This is because SCSs have



(a) Enlarged version of the single pulse in the SCS echo.

(b) Frequency spectrum of the single pulse in the SCS echo.

Figure 4.5: *Display the SCS echo with transmit frequency at* 1.83*MHz in the time and the frequency domains.*



(a) Enlarged version of the 4^{th} pulse in the MB echo.

(b) Frequency spectrum of the 4^{th} pulse in the MB echo.

Figure 4.6: *Display the MB echo with transmit frequency at* 1.83*MHz in the time and the frequency domains.*

more scattering than MBs. The corresponding spectrum is displayed in Figure 4.6(b).

Furthermore, from the computational complexity point of view, mathematical expressions of three different sinusoidal models are compared. Compared to the basic sum of sinusoidal model, the ESM introduces one more parameter, the decaying factor λ_0; the GASM also introduces an additional parameter, i.e. standard deviation of the Gaussian window α, if the center mean of the window is fixed at the midpoint of the pulse segment. Therefore, the basic sum of sinusoidal model is the simplest signal

model amongst others, and has the least computational complexity when the estimation of model parameters is performed. As a result, it is utilised as the first signal model to have a try for modeling the pulse segments for measured ultrasound echoes. The ESM and the GASM are the extensions of the basic sum of sinusoidal model, and can be considered in the future work.

4.3.4.2 Model structure selection for the full length signal

The signal models presented in Section 4.3.4.1 are only valid for a single pulse segment rather than a full length signal with multiple pulse segments in it. In other words, the proposed signal model does not take account the time domain information, i.e. the pulse location in the full length signal. In order to estimate temporal and spectral contents simultaneously, the signal model needs to be extended to incorporate the time domain information. Assume there are N data points in the signal $y(t)$, in which $t = 1, 2, \ldots, N$. In the time domain, if there are no pulses at all, the signal $y(t)$ will be a noise vector $n(t)$, as denoted in (4.7a). If there is one single pulse present, the signal model can be expressed in (4.7b). Since the location of the single pulse $s(t)$ is between T_1 and T_2, the vector of change-points, which indicate the pulse location in the signal, can be denoted by $[T_0, T_1, T_2, T_3]$, where $T_0 \triangleq 0, T_3 \triangleq N - 1$.

$$Y_0 \;:\; y(t) = n(t) \tag{4.7a}$$

$$Y_1 \;:\; y(t) = \begin{cases} n(t) & \text{if } T_0 \leq t \leq T_1 - 1 \\ s(t) + n(t) & \text{if } T_1 \leq t \leq T_2 - 1 \\ n(t) & \text{if } T_2 \leq t \leq N - 1\,, \end{cases} \tag{4.7b}$$

In Y_1, $s(t)$ represents a single pulse segment of the signal and is modeled as a sum of sinusoidal model, as proposed in Section 4.3.2. The expression of $s(t)$ is defined according to (4.3):

$$s(t) = \sum_{k=1}^{K} a_{c,k} \cos(\omega_k t) + a_{s,k} \sin(\omega_k t). \tag{4.8}$$

Nevertheless, the measured ultrasound MB responses usually have more than one pulse segment in the echo signals, unlike the single pulse SCS responses. Therefore, the signal models given in (4.7) can be further extended to a model that is able to incorporate temporal information of multiple pulse segments. Figure 4.7 illustrates a sketch of a noise-free signal $s(t)$ with m pulse segments, each of which has different

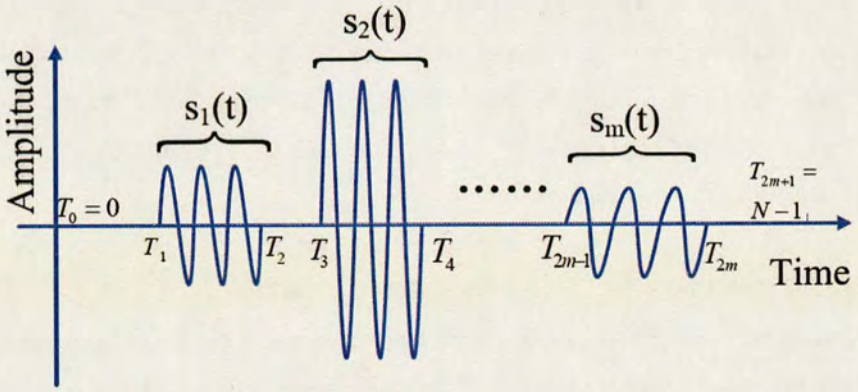


Figure 4.7: *Sketch of the description of a signal model $s(t)$.*

spectral contents and various amplitudes. The pulse locations in the signal model are denoted by a vector $[T_1, \ldots, T_{2m}]$, which represents the start and end points for all m pulses. The total length of the signal is N, and $T_0 \triangleq 0$, $T_{2m+1} \triangleq N-1$. Accordingly, the vector $[T_0, T_1, \ldots, T_{2m}, T_{2m+1}]$ represents a complete set of change-points in the signal, which indicates pulse locations in the time domain.

According to Figure 4.7, the signal model can also be defined in a mathematical way:

$$Y_0 \quad : \quad y(t) = n(t) \tag{4.9a}$$

$$Y_{k_m} \quad : \quad y(t) = \begin{cases} n(t) & \text{if } T_{2i-2} \le t \le T_{2i-1} - 1\,, \\ s_i(t) + n(t) & \text{if } T_{2i-1} \le t \le T_{2i} - 1\,, \\ n(t) & \text{if } T_{2m} \le t \le T_{2m+1}\,, \end{cases} \tag{4.9b}$$

where $i \in (1, m)$ and

$$s_i(t) = \sum_{p=1}^{k_i} a_{c_{p,k_i}} \cos(\omega_{p,k_i} t) + a_{s_{p,k_i}} \sin(\omega_{p,k_i} t). \tag{4.9c}$$

In this model, Y_0 denotes there is no pulse in the signal at all; Y_{k_m} denotes there are m pulse segments in the signal. The indices $k_1, k_2, \ldots, k_m$ denote the number of superpositioned sinusoids in m pulse segments respectively. In each ith segment, $a_{c_{p,k_i}}$, $a_{s_{p,k_i}}$ and ω_{p,k_i} are the amplitudes and the radial frequency of the pth sinusoid for the ith segment with k_i sinusoids. Moreover, $\{n(t), t = 0, \ldots, N-1\}$ is a sequence of a zero-

mean Gaussian noise. As a result, the measured ultrasound SCS echoes and MB echoes can be modeled as a signal with multiple pulses, each of which has a superpositioned sinusoids, as stated in (4.9).

4.4 Parameter estimation using Bayesian inference

Statistical inference is about drawing conclusions or making inferences about the data through quantitative measurements [99]. Probability distributions are often used to quantify uncertainties of the measurements. The specification of probability distributions to the certain measurements and some relations among them define a statistical model. Once a model is built, there are many ways to perform the inference. Bayesian inference is one popular category of statistical inference. It is able to produce a measure of the estimated parameter with a variance. In the process of Bayesian inference, inference is made by constructing the joint probability distribution of all unobserved quantities on the basis of all those that are known. In this case, Bayes's theorem should be introduced in the first place.

4.4.1 Bayes's theorem

Suppose there is a vector of observations $\mathbf{y}$. In Bayesian statistics, a set of parameters, $\mathbf{\Psi}$, denoted in terms of probability statements, is conditional on the observed value of $\mathbf{y}$ and is written as $p(\mathbf{\Psi}|\mathbf{y})$. The joint probability density function, $p(\mathbf{\Psi}, \mathbf{y})$, can be written as a product of the joint *prior* density of the parameters, $p(\mathbf{\Psi})$, and the data distribution conditioned on the parameters $p(\mathbf{y}|\mathbf{\Psi})$, which is also termed as the likelihood function. It is given by:

$$p(\mathbf{\Psi}, \mathbf{y}) = p(\mathbf{\Psi})p(\mathbf{y}|\mathbf{\Psi}) \tag{4.10}$$

According to Bayes's theorem, which is addressed in many textbooks about Bayesian data analysis [99,100], the joint distribution of parameters conditioned on the observed data sequence $\mathbf{y}$, $p(\mathbf{\Psi}|\mathbf{y})$, is termed as the joint *posterior* density, and can be calculated as:

$$p(\mathbf{\Psi}|\mathbf{y}) = \frac{p(\mathbf{\Psi})p(\mathbf{y}|\mathbf{\Psi})}{p(\mathbf{y})} \propto p(\mathbf{\Psi})p(\mathbf{y}|\mathbf{\Psi}). \tag{4.11}$$

The factor $p(\mathbf{y})$ can be omitted as a normalizing constant, except for in the case of model selection.

4.4.2 Bayesian inference

In opposition to frequency probability in the classical inference, the Bayesian approach is developed when the values of observations, $\mathbf{y}$, are initially uncertain and thus must be described using a probability distribution with density, $p(\mathbf{y}|\mathbf{\Psi})$. $\mathbf{\Psi}$ represents parameters of interest that need to be known in order to fully describe the observations. The Bayesian approach incorporates the parameter information, $\mathbf{\Psi}$, to the analysis by using a probability density, $p(\mathbf{\Psi})$, even when this information is not precise. However, the classical frequentist approach does not admit this information as it has not been observed and thus is not subject to empirical verification. There used to be much argument about these two approaches, for example in [101–103]. Nowadays, the ultimate standard is the applicability of the theory [100]. The Bayesian approach is more applicable because it includes more elements into the analysis, and the techniques described in Bayesian paradigm are able to solve more complex problems.

In Bayesian inference, there are two ingredients: the observational distribution, $p(\mathbf{y}|\mathbf{\Psi})$, and the prior distribution, $p(\mathbf{\Psi})$. The first ingredient can be represented as a function of $\mathbf{\Psi}$, which is the likelihood function, $L(\mathbf{\Psi}) \triangleq p(\mathbf{y}|\mathbf{\Psi})$. It provides the probability of realizing the observations given the values of the parameters and noise statistics under an assumed signal model. The most popular signal model in Bayesian inference is a Hierarchical Bayes model, which will be described in Section 4.4.2.1. The second ingredient is the prior density, which contains the probability distribution of $\mathbf{\Psi}$ before the observations of $\mathbf{y}$ are made. The selection of prior distributions will be discussed in 4.4.2.2. Then the inference is based on the probability distribution of $\mathbf{\Psi}$ after the observations of $\mathbf{y}$, which is called posterior distribution and denoted as $p(\mathbf{\Psi}|\mathbf{y})$, are obtained by Bayes's theorem. The posterior distribution is also termed as $\pi(\mathbf{\Psi})$, following the traditional custom in methods of inference based on stochastic simulations.

4.4.2.1 Hierarchical Bayes model

In Bayesian analysis, it is more powerful to use a rich statistical model than a simpler model. A more complicated statistical model is able to reflect a realistic problem more completely. Assume a data vector, $\mathbf{y}$, and a parameter, Ψ. Using Bayesian analysis, a posterior probability, $p(\Psi|\mathbf{y}) \propto p(\Psi)p(\mathbf{y}|\Psi)$, can be achieved by the given prior probability, $p(\Psi)$, and the likelihood, $p(\mathbf{y}|\Psi)$.

Usually, the prior $p(\Psi)$ depends on another parameter ϕ that is not included in the

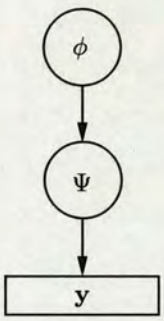


Figure 4.8: *An example of the simplest hierarchical Bayes model.*

likelihood. Then the prior will be replaced by a conditional prior, $p(\Psi|\phi)$. Therefore, the posterior distribution will be rewritten as:

$$p(\Psi, \phi|\mathbf{y}) \propto p(\mathbf{y}|\Psi)p(\Psi|\phi)p(\phi). \tag{4.12}$$

It is the simplest example of a hierarchical Bayes model. The procedure can be repeated in the exact same format. For instance, ϕ may have an additional parameter, χ, which will have its own parameters. χ is called the hyperparameter of $p(\Psi)$. In this way, there may be several stages in the specification of the joint prior distribution. Generally, the higher the stage, the harder is the specification of the distribution. However, models seldom have more than three stages. Also, the prior at the higher stage is often set to be non-informative. There are many examinations about the hierarchical structures, which can be found in [104–106].

The directed acyclic graph (DAG) is a useful graphical tool for representing a hierarchical Bayes model. In the diagram, the observational data is represented as the root of the graph; each variable is represented as a separate node pointing to the node that depends on it [99]. Figure 4.8 illustrates the simplest example of the hierarchical Bayes model in (4.12). In this figure, the observational data, $\mathbf{y}$, is denoted by a rectangular; the variables, Ψ and ϕ, are denoted by circles.

4.4.2.2 Specifications of prior distributions

There are various ways of specifying prior distributions. The non-informative prior distribution is one of the simplest choice. However, the inherent anomaly of these distributions may lead to improper distributions, which do not integrate to 1, i.e. $\int p(\mathbf{\Psi})d\mathbf{\Psi} \neq 1$. Moreover, there are many different definitions of non-informative

prior distributions. One most commonly accepted definition is Jeffrey's prior [107]. It is presented by $p(\mathbf{\Psi}) \propto |I(\mathbf{\Psi})|^{1/2}$, in which $I(\mathbf{\Psi})$[1] is the expected Fisher information matrix about $\mathbf{\Psi}$. Generally, these vague prior specifications can lead to proper posterior distributions and inference can be made easily. However, in some cases, there are exceptions and the posterior distributions will be improper. Therefore, they must be used with caution.

An alternative specification of prior distribution is the conjugate prior [100]. In Bayesian probability theory, if the posterior distribution of a parameter $p(\Psi|\mathbf{y})$ is in the same family as the prior probability distribution $p(\Psi)$, then the class of $p(\Psi)$ is said to be conjugate to the class of likelihood function $p(\mathbf{y}|\Psi)$. Moreover, the prior and the posterior are termed as conjugate distributions, and the prior is termed as a conjugate prior for the likelihood. Using conjugate priors, the posterior distribution is analytically tractable and its simplifications become easier. The calculation of posterior only involves a change in the hyperparameters. Nevertheless, since there is a restriction imposed on the form of the prior, the conjugate prior may not be adequate for one's prior for the state of uncertainty in some realistic cases. Therefore, there is always a dilemma between tractability and realism.

4.4.3 Bayesian inference on general linear models

The general linear model has a large number of applications and can be described in terms of a linear combination of basis functions with an additive Gaussian noise component [108]. The general linear model can be written in the form of:

$$\begin{aligned} \mathbf{y} &= \mathbf{s} + \mathbf{n} \\ &= \mathbf{D}\mathbf{a} + \mathbf{n}, \end{aligned} \tag{4.13}$$

where, $\mathbf{y}$, is an $N \times 1$ vector of observational data and $\mathbf{n}$ is an $N \times 1$ vector of identically independent distributed (i.i.d.) Gaussian noise. The matrix $\mathbf{D}$ has a size of $N \times Q$ and the vector of linear coefficients, $\mathbf{a}$, has a size of $Q \times 1$. Each element of the $Q \times 1$ vector is a linear coefficient corresponding to a particular column of the matrix $\mathbf{D}$. Moreover, the matrix $\mathbf{D}$ contains all the signal information, such as spectral and temporal contents.

[1] $I(\mathbf{\Psi}) \triangleq -E\left[\frac{\partial^2}{\partial \mathbf{\Psi}^2} \ln L(\mathbf{\Psi})|\mathbf{\Psi}\right]$, where $L(\mathbf{\Psi})$ is the likelihood function of $\mathbf{\Psi}$. $E[\mathbf{X}|\mathbf{Y}]$ is defined as the conditional expectation.

The likelihood is the probability of realizing the data given the value of parameters, the signal model and the noise statistics [108]. It was shown in [109] that the likelihood function is equivalent to the joint density of the residuals when the noise is i.i.d.. In this case, the likelihood function for the general linear model can be given by:

$$\begin{aligned} p(\mathbf{y}|\mathbf{\Psi}) &= (2\pi\sigma_n^2)^{(-\frac{N}{2})} \exp\left[-\frac{\mathbf{n}^T\mathbf{n}}{2\sigma_n^2}\right] \\ &= (2\pi\sigma_n^2)^{(-\frac{N}{2})} \exp\left[-\frac{(\mathbf{y}-\mathbf{Da})^T(\mathbf{y}-\mathbf{Da})}{2\sigma_n^2}\right], \end{aligned} \tag{4.14}$$

where σ_n^2 is the variance of the Gaussian noise, and $\mathbf{\Psi}$ is a vector of model parameters in matrix $\mathbf{D}$.

When choosing prior distributions for model parameters in Bayesian inference, one important criterion is convenience. Therefore, in the case of a general linear model, uniform priors are assumed over each element in the coefficients vector, $\mathbf{a}$, and Jeffreys's prior is assigned to the variance of the Gaussian noise, σ_n^2. After the integration of the linear amplitude coefficients, $\mathbf{a}$, and the noise variance, σ_n^2, as derived in [108], the posterior probability density for the model parameters can be simplified as:

$$p(\mathbf{\Psi}|\mathbf{y}) \propto \frac{[\mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{D}(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{y}]^{(Q-N)/2}}{\sqrt{\det(\mathbf{D}^T\mathbf{D})}}. \tag{4.15}$$

In the following subsections, there are three examples of general linear models to show how Bayesian inference is used for parameter estimation. Section 4.4.3.1 studies a single frequency signal model and Section 4.4.3.2 examines a signal with two frequency components. They present a more detailed description of Bayesian inference with its application to spectral analysis, usually termed as Bayesian spectral estimation. It is shown that Bayesian inference explores a new perspective of spectral estimation. Jaynes [96] first applied the principle of Bayesian inference into spectral analysis and estimation in 1987. In a Bayesian analysis, the width of a spectral peak, which reflects the accuracy of the frequency estimate, is determined by the duration of the observed data, the S/N [2] and the number of data points. Gregory [31] also presented an approximation equation of calculating the frequency resolution in Hz if $S/N > 1$:

$$\triangle f \approx \left\{1.6 \cdot S/N \cdot T_s \cdot \sqrt{N}\right\}^{-1}, \tag{4.16}$$

[2]The S/N is defined as the ratio of the root mean square signal amplitude to the noise standard deviation.

where T_s denotes the data duration in seconds and N denotes the number of data points in T_s. In addition to the spectral content, the temporal content in the time domain can also be estimated using Bayesian inference. It is able to incorporate more parameters of interests in the proposed signal model, not only the frequency components. For example, Section 4.4.3.3 examines a signal with one frequency and one change-point.

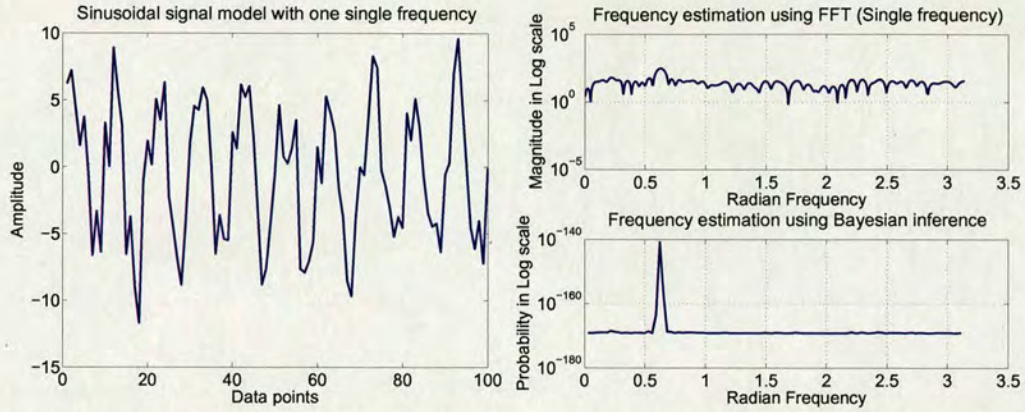
4.4.3.1 One frequency signal model

In this example, a signal with a single frequency component, ω_1, is simulated with a length of $N = 100$ data points. It is the simplest case of the sum of sinusoidal signal model. For this special case, the matrix, $\mathbf{D}$, and the vector of coefficients, $\mathbf{a}$, in (4.13) can be specified as:

$$\mathbf{D} = \mathbf{D}(\omega_1) \triangleq \begin{bmatrix} \cos(\omega_1 t_1) & \sin(\omega_1 t_1) \\ \cos(\omega_1 t_2) & \sin(\omega_1 t_2) \\ \vdots & \vdots \\ \cos(\omega_1 t_N) & \sin(\omega_1 t_N) \end{bmatrix},$$

and $\mathbf{a} \triangleq [a_{c,1}, a_{s,1}]^T$ respectively.

The parameter setting for this simulated single frequency signal is given in Table 4.2. Moreover, Figure 4.9(a) displays the simulated signal in the time domain with a SNR of 5dB. The bottom panel in Figure 4.9(b) shows the probability of the estimated frequency. It is calculated using the general format of the posterior distribution in (4.15). As ω_1 is the only parameter of interest in the matrix $\mathbf{D}$, $\mathbf{\Psi}$ is replaced by ω_1. The prominent peak in the figure indicates the most probable angular frequency in the signal model. It is compared to the corresponding frequency spectrum based on the DFT, which is shown in the upper panel in Figure 4.9(b). A well known property of the FT-based frequency estimation is that the width of any spectral peak depends only on the duration of the signal and not on the signal to noise level. By contrast, in Bayesian analysis, the width of spectral peak is determined by many factors, e.g. signal to noise level, duration of the signal and the number of data points in that duration, which is given in (4.16). Therefore, in the case of low signal to noise level, the frequency resolution is better if the duration of a signal is fixed. In Figure 4.9(a), the SNR is 5dB and the signal to noise level according to the definition in (4.16) is 1.7783. It is demonstrated



(a) Simulated signal with single frequency in the time domain.

(b) FFT spectrum and Bayesian inference of the simulated signal with single frequency.

Figure 4.9: *A signal with a single frequency and its frequency estimation.*

k	$E_k \triangleq a_{s,k}^2 + a_{c,k}^2$	$-\arg\tan(a_{s,k}/a_{c,k})$	ω_k
1	40	$\pi/3$	0.2π

Table 4.2: *Parameter settings for the signal model with a single frequency*

in Figure 4.9(b) that the spurious noise features are suppressed and the width of the peak is much narrower than the peak in the DFT-based frequency spectrum.

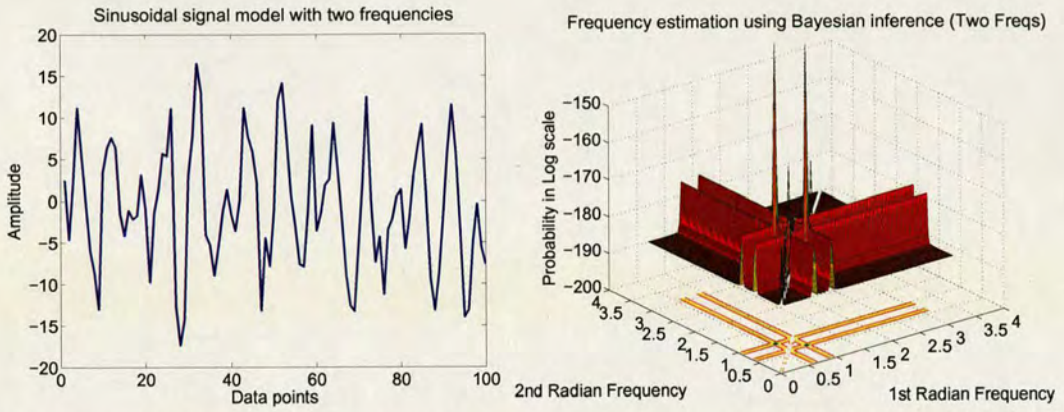
4.4.3.2 Two frequency signal model

The second example has two frequency components, ω_1 and ω_2 in the signal, and has a length of $N = 100$ data points. The signal can be modeled using (4.13) with a different definition of the matrix, $\mathbf{D}$, and the coefficients vector, $\mathbf{a}$, which are expressed as:

$$\mathbf{D} = \mathbf{D}(\omega_1, \omega_2) \triangleq \begin{bmatrix} \cos(\omega_1 t_1) & \sin(\omega_1 t_1) & \cos(\omega_2 t_1) & \sin(\omega_2 t_1) \\ \cos(\omega_1 t_2) & \sin(\omega_1 t_2) & \cos(\omega_2 t_2) & \sin(\omega_2 t_2) \\ \vdots & \vdots & \vdots & \vdots \\ \cos(\omega_1 t_N) & \sin(\omega_1 t_N) & \cos(\omega_2 t_N) & \sin(\omega_2 t_N) \end{bmatrix},$$

and $\mathbf{a} \triangleq [a_{c,1}, a_{s,1}, a_{c,2}, a_{s,2}]^T$.

The parameter setting for this simulated signal with two frequency components is given in Table 4.3. Moreover, Figure 4.10(a) displays the simulated signal in the time domain with a SNR of 5dB. By replacing $\mathbf{\Psi}$ by (ω_1, ω_2) in (4.15), the joint posterior probability of the two frequencies ω_1 and ω_2 can be obtained. Unfortunately, this



(a) Simulated signal with two frequencies in the time domain.

(b) Surface and contour plot for the joint probability of two estimated frequencies.

Figure 4.10: *Bayesian inference of the simulated signal with two frequencies.*

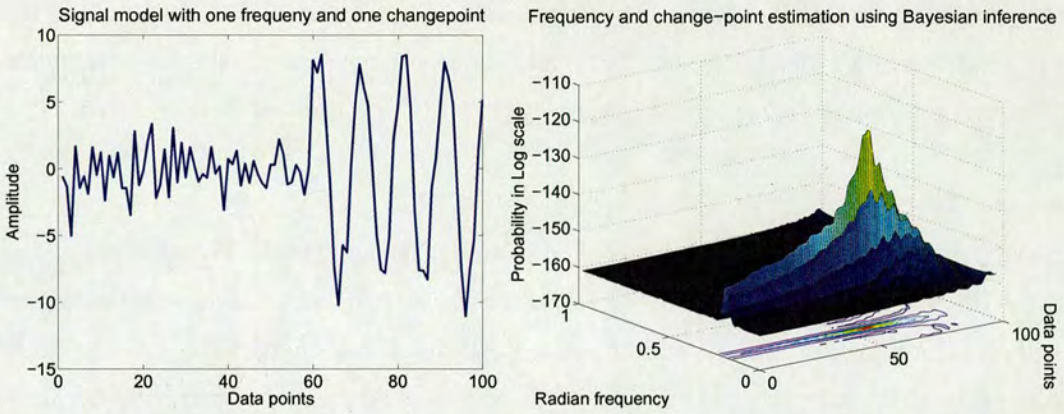
k	$E_k \triangleq a_{s,k}^2 + a_{c,k}^2$	$-\arg\tan(a_{s,k}/a_{c,k})$	ω_k
1	40	$\pi/3$	0.2π
2	40	$\pi/4$	0.3π

Table 4.3: *Parameter settings for the signal model with two frequencies*

obtained joint posterior probability is in a two-dimensional space, which should be plotted in a 3D figure. Figure 4.10(b) illustrates a 3D surface plot and a contour plot of the joint distribution of two angular frequencies. The values of two frequency components are determined by the peak locations. The symmetry of Figure 4.10(b) indicates the equal probability of the two frequencies, ω_1 and ω_2, which are interchangeable. If $\omega_1 = \omega_2$, $\mathbf{D}^T\mathbf{D}$ becomes singular and thus the surface plot of the joint posterior distribution will have no peaks at all.

4.4.3.3 One frequency and one change-point signal model

Previous two examples studied in Section 4.4.3.1 and Section 4.4.3.2 are two cases of the sum of sinusoidal signal model specially designed for spectral estimation. The joint posterior distribution obtained only contains the frequency information of the signal model. In this section, a signal which has one change-point, T_1, and one frequency component, ω_1, is proposed. The change-point is a non-linear parameter in the matrix of basis function $\mathbf{D}$. Before the change-point, the signal has Gaussian noise only; after the change-point, the signal is a single frequency sinusoid embedded in Gaussian



(a) Simulated signal with one frequencies and one change-point in the time domain.

(b) Surface and contour plot for the joint probability of the estimated frequency and change-point.

Figure 4.11: *Bayesian inference of the simulated signal with one frequency and one change-point.*

Parameters	$E_1 \triangleq a_{s,1}^2 + a_{c,1}^2$	$-\arg\tan(a_{s,1}/a_{c,1})$	ω_1	T_1
Assigned values	40	$\pi/3$	0.2π	60

Table 4.4: *Parameter settings for the signal model with one frequency and one change-point*

noise. The joint posterior probability of parameters, $p(\omega_1, T_1|\mathbf{y})$ can also be calculated using (4.15). The differences are the definitions of basis function matrix, $\mathbf{D}$, and the linear coefficients, $\mathbf{a}$, which can be defined as:

$$\mathbf{D} = \mathbf{D}(\omega_1, T_1) \triangleq \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \cos(\omega_1 t_{T_1}) & \sin(\omega_1 t_{T_1}) \\ \cos(\omega_1 t_{T_1+1}) & \sin(\omega_1 t_{T_1+1}) \\ \vdots & \vdots \\ \cos(\omega_1 t_N) & \sin(\omega_1 t_N) \end{bmatrix},$$

and $\mathbf{a} \triangleq [a_{c,1}, a_{s,1}]^T$ respectively.

The parameter setting for this simulated signal with one frequency and one change-point is given in Table 4.4. Figure 4.11(a) displays this simulated signal in the time domain. The SNR of it is also set to 5dB. The change in data points can be clearly seen

at $T_1 = 60$. Figure 4.11(b) shows a surface plot and a contour plot of the joint posterior distribution of the change-point, T_1, and the frequency component, ω_1. The peak location determines the estimated frequency, $\hat{\omega}_1$ and the estimated change-point, $\hat{T}_1$.

4.4.3.4 Discussion

Having obtained the probabilities of signal model parameters in Figure 4.9(b), Figure 4.10(b) and Figure 4.11(b), the optimisation techniques for finding the peaks are still required in order to determine the exact estimated parameter values. The examples addressed above have at most two parameters to be estimated and thus the joint posterior distribution of these parameters can be represented directly either in a 2D plot or a 3D surface plot. Due to low variances of the peaks in figures, determining the peak locations is not difficult if a 2D plot or a 3D plot is available.

Unfortunately, in many real application cases, there are usually more than two parameters that need to be estimated, for instance, there are three frequency components, or there are two change-points and one frequency component in a signal, the parameters are difficult to estimate with the given joint posterior distribution. The only way to find the marginal probability for a specific parameter is to integrate out all the nuisance parameters. When more and more parameters are involved, the marginal posterior distribution for some parameter cannot be easily obtained in an analytic form and the numerical calculation of the integral can be very complicated and difficult. Accordingly, approximate Bayesian computation is required to tackle the computationally intractable problems. One of the most popular approximation approaches in Bayesian inference is the MCMC technique. A detailed review of MCMC algorithms will be presented in Chapter 5.

Furthermore, using MCMC algorithms, Bayesian inference can easily be extended to incorporate model order selection into spectral analysis with an employed reversible jump technique. In other words, the number of frequency in a signal does not need to be known in advance and thus can be regarded as an extra parameter to be estimated. In addition to solving model order selection problem, Bayesian inference is also able to incorporate the time domain information of a signal into the estimation and thus provides the estimated change-points in the signal simultaneously. As a result, all the parameters of interest including temporal content in the time domain and spectral content in the frequency domain, can be estimated at the same time within a whole single Bayesian framework.

4.5 Bayesian inference for the proposed signal model

In Section 4.3.4.2, the proposed signal model for multiple pulse echo signals is defined in (4.9). Once the signal model is determined, a *parametric* estimation approach using Bayesian inference can be applied, taking account of all the parameters of interest. As the inference will be approximated using MCMC algorithms, which are difficult to examine for the proposed sophisticated signal model, the detailed description about the estimation algorithm will be presented in Chapter 6. This section only gives a brief introduction of the proposed signal model for measured ultrasound echo signals, and presents the corresponding likelihood functions for the specific model.

4.5.1 Likelihood of a full length signal model

According to the proposed signal model in (4.9), a general form of the signal model can be written in a matrix-vector form:

$$\mathbf{y} = \mathbf{D}(\boldsymbol{\omega}_{k_m}, \boldsymbol{T}_{2m})\mathbf{a}_{k_m} + \mathbf{n}, \tag{4.17a}$$

where $\mathbf{a}_{k_m} \triangleq [\boldsymbol{a}_{k_1}, \boldsymbol{a}_{k_2}, \ldots, \boldsymbol{a}_{k_m}]^T$, in which $\boldsymbol{a}_{k_i} \triangleq [(a_{c_1,k_i}, a_{s_1,k_i}), \ldots, (a_{c_{k_i},k_i}, a_{s_{k_i},k_i})]^T$, $(i = 1, \ldots, m)$ represents the amplitudes of frequency components in each pulse segment. Moreover, $\mathbf{D}(\boldsymbol{\omega}_{k_m}, \mathbf{T}_{2m})$ can be defined as follows:

$$\mathbf{D}(\boldsymbol{\omega}_{k_m}, \mathbf{T}_{2m}) \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \boldsymbol{D}_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{D}_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \boldsymbol{D}_m \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{4.17b}$$

The matrix $\mathbf{D}(\boldsymbol{\omega}_{k_m}, \mathbf{T}_{2m})$ has a size of $N \times \sum_{i=1}^{m} 2k_i$, where N is the full length of the observations, $\mathbf{y}$, and k_i represents the number of frequencies in each pulse segment i. It contains the information about change-points $[T_1, T_2, \ldots, T_{2m}]$ and spectral contents $[\boldsymbol{\omega}_{k_1}, \boldsymbol{\omega}_{k_2}, \ldots, \boldsymbol{\omega}_{k_m}]$ for m different pulse segments, which are related to temporal and

spectral contents in the full length signal respectively. $\boldsymbol{D}_i$ in the matrix $\mathbf{D}(\boldsymbol{\omega}_{k_m}, \mathbf{T}_{2m})$ represents the ith pulse segment in the signal, which has its own parameters. It can be defined as:

$$\boldsymbol{D}_i \triangleq \begin{bmatrix} B(\omega_{k_1} t(T_{2i-1})) & \dots & B(\omega_{k_m} t(T_{2i-1})) \\ B(\omega_{k_1} t(T_{2i-1}+1)) & \dots & B(\omega_{k_m} t(T_{2i-1}+1)) \\ \vdots & \ddots & \vdots \\ B(\omega_{k_1} t(T_{2i}-1)) & \dots & B(\omega_{k_m} t(T_{2i}-1)) \end{bmatrix}, \tag{4.17c}$$

where $B(\cdot) \triangleq [\cos(\cdot), \sin(\cdot)]$ and it operates on an element by element basis. The size of $\boldsymbol{D}_i$ is $\{(T_{2i} - T_{2i-1}) \times 2k_i\}$, in which T_{2i-1} and T_{2i} correspond to the start and end points for the ith pulse segment. In each $\boldsymbol{D}_i$, the number of sinusoids k_i, indicating different number of frequencies in the ith pulse segment, is unknown. Other parameters conditioned on k_i in each single pulse segment are also unknown. Moreover, the noise term is assumed as a zero-mean white Gaussian noise, which is defined as: $\mathbf{n} \triangleq [n(0), n(1), \cdots, n(N-1)]^T$.

In this signal model, the estimation for pulse locations, $\mathbf{T}_{2m}$, is performed using a full length data. As the noise is an i.i.d. Gaussian noise with variance σ_n^2, the corresponding likelihood function is expressed as:

$$p(\mathbf{y}|\boldsymbol{k}_m, \boldsymbol{\Psi}_m, \boldsymbol{T}_{2m}) = (2\pi\sigma_n^2)^{-N/2} \times \exp\left\{-\frac{\| \mathbf{y} - \mathbf{D}(\boldsymbol{\omega}_{k_m}, \mathbf{T}_{2m})\mathbf{a}_{k_m} \|^2}{2\sigma_n^2}\right\}, \tag{4.18}$$

where $\|A\|^2 \triangleq A^T A$, and N is the full length of signal model. $\boldsymbol{\Psi}_m \triangleq (\boldsymbol{\omega}_{k_m}, \mathbf{a}_{k_m}, \sigma_n^2)$ contains the parameters in the full length signal model.

4.5.2 Likelihood of a single pulse segment

Each single pulse segment in the signal model has its own parameters that are used to fit each block of data $\mathbf{y}_i$, which should be examined individually. Therefore, for the ith pulse segment with a length of $N_i = T_{2i} - T_{2i-1}$, $(i = 1, \dots, m)$, $\mathbf{y}_i$ is defined as:

$$\mathbf{y}_i = \boldsymbol{D}_i \boldsymbol{a}_{k_i} + \mathbf{n}_i. \tag{4.19}$$

Based on this model of a single pulse segment, the likelihood function for an individual data block $\mathbf{y}_i$ is presented as:

$$p(\mathbf{y}_i|k_i, \boldsymbol{\theta}_{k_i}, T_{2i}, T_{2i-1}) = (2\pi\sigma_{k_i}^2)^{-N_i/2} \times \exp\left\{-\frac{\| \mathbf{y}_i - \boldsymbol{D}_i(\boldsymbol{\omega}_{k_i}, T_{2i-1}, T_{2i})\boldsymbol{a}_{k_i} \|^2}{2\sigma_{k_i}^2}\right\}, \tag{4.20}$$

where $\boldsymbol{\theta}_{k_i} \triangleq (\boldsymbol{\omega}_{k_i}, \boldsymbol{a}_{k_i}, \sigma_{k_i}^2)$ contains the parameters in the ith pulse segment.

As a result, the *parametric* estimation is divided into two steps in the estimation algorithm. The first step is estimation for pulse locations, $\mathbf{T}_{2m}$, which is performed in a full length signal $\mathbf{y}$. The second step is estimation for spectral content, which is performed in a particular pulse segment $\mathbf{y}_i$, and the same procedure of estimation is performed for all m pulses individually.

4.6 Model order selection

In *parametric* estimation methods, selection of several integer valued parameters is equally important for the specification of a signal model. These integer valued parameters of the model can be orders of an ARMA model, or the number of sinusoidal components in a noisy signal. In either of these cases, the integer valued parameters determine the dimension of parameter vector of the signal model, which is often termed as model order.

4.6.1 Traditional selection criteria

Most of the model order selection techniques are based on an information criterion, which has been widely used in time series analysis to monitor the modeling error and to select the order at which this error enters a steady state. The information criterion is able to provide a measure of information between the true distribution of data and the distribution of a specific model, in terms of the model order. The Kullback-Leibler (KL) information [110] is often used to measure the difference of two probability distributions. It can provide the distance between the true distribution of the observed data and the theoretical model, and it is measured by the log-likelihood function of the distributions. The derivation of the application of the KL information on model order selection can be found in [110]. Several criteria are proposed on the basis of the KL information and aim at minimizing the residual sum of squares.

Some well-known model order selection criteria are briefly introduced in the fol-

lowing. They are all based on the maximum likelihood (ML) estimate of the variance of the residuals between the proposed model and the data samples. Generally, ML estimator is a statistical method for fitting a signal model to the measured data samples and providing estimates for model parameters [111]. It selects the model parameter values that make the signal model more likely to the measured data. There are a variety of other estimators in the literature. A brief description will be presented in Section 4.7.

Akaike [112] in 1969 proposed an order selection criterion, final prediction error (FPE), which is defined as:

$$\mathrm{FPE}(k) = \frac{N+k}{N-k}\hat{\sigma}_k^2, \tag{4.21}$$

where N is the number of data samples, and $\hat{\sigma}_k^2$ is the modeling error, which is the ML estimate of the variance of the residuals of the estimated autoregressive (AR) model with the kth order. It works well for the AR process [113]. In 1974, Akaike [114] introduced a new criterion called the Akaike information criterion (AIC). When there are k independent adjusted parameters fitted to data, it can be defined in general as:

$$\mathrm{AIC}(k) = -2\ln(\text{maximum likelihood}) + 2k, \tag{4.22}$$

where ln denotes the natural logarithm operation. Specifically for an AR model with model order k, this criterion in (4.22) becomes:

$$\mathrm{AIC}_{(AR)}(k) = -2\ln p_k(\mathbf{y}, \hat{\mathbf{\Psi}}_k) + 2k. \tag{4.23}$$

In the above equation, $\mathbf{y}$ is the assumed AR process, and $\mathbf{\Psi}_k$ is the parameter vector with kth order in the model. Its likelihood function is denoted as : $p_k(\mathbf{y}, \mathbf{\Psi}_k)$. $\hat{\mathbf{\Psi}}_k$ is the ML estimate of $\mathbf{\Psi}_k$, which is given by:

$$\hat{\mathbf{\Psi}}_k = \arg\max_{\mathbf{\Psi}_k} \ln p_k(\mathbf{y}, \mathbf{\Psi}_k). \tag{4.24}$$

Moreover, it is assumed that the data is embedded in real-valued i.i.d. Gaussian white-noise with mean zero and a variance of σ_k^2. After a simple derivation, which can be found in [110], the first term in (4.23) can be simplified as:

$$-2\ln p_k(\mathbf{y}, \hat{\mathbf{\Psi}}_k) = N\ln\hat{\sigma}_k^2 + \text{constant}, \tag{4.25}$$

where $\hat{\sigma}_k^2$ has the same definition as used in (4.21). Furthermore, the AIC can be generalized to the following criterion and is termed the generalized information criterion (GIC) [115]:

$$\mathrm{GIC}(k) = N \ln(\hat{\sigma}_k^2) + \alpha k, \tag{4.26}$$

where α is a positive constant. With $\alpha = 2$, the GIC is equivalent to the AIC. The FPE criterion can also be viewed as a special case of the extended GIC criterion since the AIC and the FPE are closely related. The derivation and the investigation of their relationship can be found in [116]. In 1978, Schwarz [117] developed a Bayesian extension of the minimum AIC procedure called Bayesian information criterion (BIC) as follows:

$$\mathrm{BIC}(k) = N \ln(\hat{\sigma}_k^2) + k \ln N. \tag{4.27}$$

The same rule in (4.27) can be obtained by a different approach based on the minimum description length (MDL) principle, which was proposed by Risannen in 1978 [118]. In 1979, an order selection criterion proposed by Hannan and Quinn [119] had a similar format as that proposed by Akaike:

$$\mathrm{HQ}(k) = N \ln(\hat{\sigma}_k^2) + kC \ln(C \ln N), \tag{4.28}$$

where C is a constant and $C > 2$. It is pointed out in [119] that the Hannan-Quinn (HQ) criterion has a good performance if the number of observations, N, is large. Another popular choice for model order selection is Parzen's criterion for autoregressive transfer function (CAT) [120], which can be denoted as:

$$\mathrm{CAT}(k) = \frac{1}{N} \sum_{i=1}^{k} \frac{N-i}{N\hat{\sigma}_i^2} - \frac{N-k}{N\hat{\sigma}_k^2}. \tag{4.29}$$

It is asymptotically equivalent to the AIC related criteria. Nevertheless, it is pointed out in [121] that for harmonic processes in noise, AIC tends to underestimate the order if the SNR is high. Furthermore, in many practical applications, these criteria do not exhibit a clear minimum which makes order selection process difficult.

4.6.2 Bayesian model selection

In addition to these model order selection criteria that are based on the information criteria, Bayesian model determination is an alternative method available for model

selection and model order selection [99]. It is based on Bayesian inference, which has already been introduced in detail in Section 4.4. It uses Bayesian model comparison, which is based on Bayes factors, to determine which model is the most appropriate one. In the Bayesian model determination approach, the dimensionality of the parameter vector is not fixed.

Suppose that a vector of observations, $\mathbf{y}$, has a collection of candidate models $\{\mathcal{M}_k, k \in \mathcal{K} = [0, 1, 2, \ldots]\}$, where k is the model order that needs to be estimated. For each given k, there is a corresponding parameter vector $\mathbf{\Psi}_k$. Take a model selection problem between two models, $\mathcal{M}_1$ and $\mathcal{M}_2$, as an example. On the basis of the observation vector, $\mathbf{y}$, the two models, $\mathcal{M}_1$ and $\mathcal{M}_2$, can be assessed by a Bayes factor:

$$K_B(\mathcal{M}_1, \mathcal{M}_2) = \frac{p(\mathbf{y}|\mathcal{M}_2)}{p(\mathbf{y}|\mathcal{M}_1)} = \frac{\int p(\mathbf{\Psi}_2|\mathcal{M}_2)p(\mathbf{y}|\mathbf{\Psi}_2, \mathcal{M}_2)d\mathbf{\Psi}_2}{\int p(\mathbf{\Psi}_1|\mathcal{M}_1)p(\mathbf{y}|\mathbf{\Psi}_1, \mathcal{M}_1)d\mathbf{\Psi}_1}, \tag{4.30}$$

where $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$ represent model parameter vectors for model $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively. This Bayes factor can be used to calculate the ratio of posterior probabilities of models $\mathcal{M}_1$ and $\mathcal{M}_2$. This ratio is often termed as posterior odds:

$$\frac{p(\mathcal{M}_2|\mathbf{y})}{p(\mathcal{M}_1|\mathbf{y})} = \frac{p(\mathcal{M}_2)}{p(\mathcal{M}_1)} \times K_B(\mathcal{M}_1, \mathcal{M}_2). \tag{4.31}$$

According to the obtained posterior odds, the selection of a single model $\mathcal{M}_1$ or model $\mathcal{M}_2$ can be determined easily. In this way, among all $\mathcal{M}_k$ models, the most appropriate model can be selected by calculating the posterior odds. However, this model selection method needs to store the information for all $\mathcal{M}_k$ models and thus is inefficient.

Furthermore, for a model where an explicit expression of the likelihood is not available or cannot be evaluated numerically, approximate Bayesian computation can be used for model selection in a Bayesian framework. Some standard computational approaches, such as MCMC methods, are available to tackle the computationally intractable problems. In particular, a reversible jump technique proposed in [122] can be combined with the MCMC algorithms to solve the model order selection problem from the applied statistics point of view within a Bayesian framework, which will be investigated in detail in Section 5.3.5. This technique is based on the reversible samplers that can jump between parameter subspaces of differing dimensionality. Therefore, it can link the model order selection technique to the parameter estimation technique, which can simplify these two steps in the process of signal model building.

4.7 Estimators

When using Bayesian inference for the estimation of each model parameter, there are a large number of data samples available from the corresponding posterior distribution. An estimator is a function of the observed data, which is used to estimate an unknown population parameter in statistics. The ML estimator has been employed in conventional model order selection criteria and has been briefly introduced in Section 4.6.1. There are a variety of estimators in the estimation literature, and a general discussion about different estimators has been given by Kay in [111]. This section introduces two widely used estimators: Maximum a posterior estimation and Minimum mean square error estimation.

4.7.1 Maximum a posteriori estimation

Maximum a posteriori estimation is considered in Bayesian statistics. A maximum a posteriori (MAP) estimate is a mode of a posterior distribution. Specifically, once N samples from a posterior distribution, $\pi(\Psi)$, are available, a histogram can be plotted to provide an estimate of the probability density. Then the MAP estimate is generated by finding the maximum of this histogram:

$$\hat{\Psi}_{MAP} = \underset{\Psi^{(i)};i=1,\ldots,N}{\arg\max}\ \pi(\Psi). \tag{4.32}$$

4.7.2 Minimum mean square error estimator

In statistics, a Minimum mean square error estimator describes a method which minimises the MSE, which is a common measure of an estimator quality. As an alternative way to estimate the parameter with the simulated N samples, the minimum mean squared error (MMSE) estimator is defined as the conditional expectation of the parameters given the obtained samples, which is given by:

$$\hat{\Psi}_{MMSE} = E\left[\pi(\Psi)|\Psi\right]. \tag{4.33}$$

4.8 Chapter summary

Based on the *parametric* estimation methods, this chapter has first addressed a general description of signal modeling and parameter estimation problems. Then according to physical understandings of measured ultrasound echo signals, a signal model has been

proposed. The proposed signal model not only contains spectral contents in multiple pulse segments, but incorporates temporal content of the full length signal as well.

After the model structure selection is discussed, Bayesian inference is employed to estimate model parameters, with applications to some general linear models. Nevertheless, the proposed signal model is too complicated and thus the posterior distribution for model parameters cannot be obtained analytically. This chapter only presents simple description of how Bayesian inference is used for the full length signal model with multiple pulses. Based on approximate Bayesian computation methods, which will be reviewed in detail in Chapter 5, Chapter 6 will provide a solution to sampling from the sophisticated posterior distributions. In this way, the temporal and spectral estimation for measured ultrasound echo signals can be achieved.

Chapter 5

Approximation approaches for Bayesian inference

Chapter 4 has proposed a signal model for measured ultrasound echo signals and employed Bayesian inference to estimate the model parameters. However, the posterior distribution of the desired parameters cannot be obtained in analytic forms. Therefore, approximate Bayesian computation methods presented in this chapter are used as sampling strategies from these posterior distributions. Moreover, MCMC algorithms are reviewed as one of the most popular standard computational approaches used in Bayesian inference.

5.1 Introduction

As discussed in Section 4.4.3.4, in many real applications, a closed form of the posterior distribution of a desired model parameter within a Bayesian framework is not available. Therefore, approximate Bayesian computation methods are required. They are simulation techniques that operate on a population of data samples and can provide a good approximation to the true posterior. In this way, approximate Bayesian computation can be used to tackle the computationally intractable problems.

In this chapter, Section 5.2 provides a brief introduction of approximation methods for Bayesian inference. In particular, commonly used stochastic simulation methods are addressed. As one of the most popular approximate Bayesian computational approaches, a detailed review of MCMC algorithms are presented in Section 5.3. It covers from the concepts of Markov chains, to some popular algorithms, such as Metropolis-Hasting (MH) algorithm and Gibbs sampler. Moreover, a reversible jump technique, which is brought into MCMC algorithm to solve the model order selection problem, is also presented.

5.2 Approximations for Bayesian inference

When a posterior distribution is analytically intractable in Bayesian inference, an inevitable way is to seek approximations of the distribution, which include deterministic and stochastic methods. Some of the deterministic methods are based on analytical approximations whereas others are based on numerical approximations. The most commonly used approximations are normal approximations and standard Laplace approximations. Detailed descriptions can be found in [100,123,124]. Nevertheless, with the increase of the dimension of the parametric space, the complexity of the deterministic techniques increases substantially. Therefore, the use of stochastic simulation for inference from the posterior distribution becomes desirable.

5.2.1 Stochastic simulation

In the real world, simulation refers to dealing with a real problem through reproduction in a controlled environment [100]. The environment can be either a computational equipment or a smaller scale reproduction system. In some systems, it has all its components known or at least deductible and thereby its behaviour has an intrinsic deterministic character. In some other systems, some of the components are subject to random fluctuations, thus can only be described by probabilistic statements rather than an exact mathematical rule. Stochastic simulation is about dealing with simulation of these systems.

From a statistical point of view, the systems mentioned above can be regarded as a complicated function of random variables. The aim is to reproduce these random variables in an environment under control regardless of the complexity. Based on the stochastic simulation, extracting relevant information in a given parent distribution is made possible. The approach, termed as Monte Carlo simulation, uses a collection of the simulated values from the distribution of interest to form a set of samples and defines a discrete distribution concentrated on the sample values. Then this distribution is an approximation to its parent distribution used for the simulation. Moreover, a histogram of these samples can be plotted. If an adequately large number of simulated values are available, the resulting histogram can be regarded as a very close approximation to the density of the distribution of interest. This is because stochastic simulation techniques have strong support in probability results, such as the law of large numbers [125], which ensures the approximation becomes increasingly bet-

ter as the number of simulated values increases. Before the stochastic simulation is addressed, some basic, direct simulation operations are presented.

5.2.2 The Monte Carlo principle

The aim of Monte Carlo simulation methods is drawing i.i.d. samples from a target distribution, $\pi(\mathbf{\Psi})$, which is often in a high dimensional space. These samples are used to form a histogram of the target distribution in order to obtain a numerical approximation of it.

Consider a function $f(\mathbf{\Psi})$ with N samples. The target probability distribution function is $\pi(\mathbf{\Psi})$ and the expected value can be defined as:

$$I_f \triangleq E\left[f(\mathbf{\Psi})\right] = \int \pi(\mathbf{\Psi}) f(\mathbf{\Psi}) d\mathbf{\Psi}. \tag{5.1}$$

As $\pi(\mathbf{\Psi})$ is difficult to evaluate analytically, using the Monte Carlo numerical approximation is an alternative:

$$\hat{\pi}(\mathbf{\Psi}) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\Psi}(\mathbf{\Psi}^{(i)}), \tag{5.2}$$

where $\delta_{\Psi}(\mathbf{\Psi}^{(i)})$ denotes an indicator function defined as:

$$\delta_{\Psi}(\mathbf{\Psi}^{(i)}) = \begin{cases} 1, & \text{if} \quad \delta_{\Psi}(\mathbf{\Psi}^{(i)}) \quad \text{locates within a small finite region around} \quad \mathbf{\Psi}^{(i)}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

Therefore, the integral in (5.1) can be approximated as follows:

$$\hat{I}_f = \int \hat{\pi}(\mathbf{\Psi}) f(\mathbf{\Psi}) d\mathbf{\Psi} = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)}). \tag{5.4}$$

According to the strong law of large numbers, i.e. the sample average converges almost surely to the expected value, the estimates $\hat{I}_f$ converges to I_f with $N \to +\infty$, as denoted in (5.5), if the samples $\mathbf{\Psi}$ are statistically independent. Furthermore, the convergence in (5.5) allows the existence of central limit theorem for $\hat{I}_f$ in (5.6), where $\sigma_f^2 \triangleq E[f^2(\mathbf{\Psi})] - I_f^2 < \infty$. In (5.5) and (5.6), $\to$ denotes the convergence in distribution.

$$\lim_{N \to +\infty} \hat{I}_f = I_f, \tag{5.5}$$

$$\lim_{N \to +\infty} \sqrt{N}(\hat{I}_f - I_f) \to \mathcal{N}(0, \sigma_f^2). \tag{5.6}$$

Algorithm 5.1: Accept-Reject Algorithm

```
for i = 1 to N do
    Sample Ψ^(i) ~ q(Ψ) and u ~ U_[0,1] ;
    if u < π(Ψ^(i)) / (M q(Ψ^(i))) then
        Accept the sample ;
    else
        Reject the sample ;
    end
end
```

If $\pi(\mathbf{\Psi})$ is a standard distribution, e.g. Gaussian, Gamma distributions, samples can be drawn straightforwardly. However, when it is not the case, generating i.i.d. samples from an arbitrary multivariate non-standard probability density function becomes intractable. More sophisticated techniques, which are based on direct sampling, importance sampling and MCMC algorithms are needed.

5.2.3 Direct sampling

Assume a complex distribution $\pi(\mathbf{\Psi})$ needs to be sampled. $q(\mathbf{\Psi})$ is another distribution which is easy to sample. These two distributions satisfy the following relationship with a proportionality constant M:

$$\pi(\mathbf{\Psi}) \leq Mq(\mathbf{\Psi}) \quad M < +\infty. \tag{5.7}$$

$q(\mathbf{\Psi})$ can be referred to as an envelope of $\pi(\mathbf{\Psi})$, and M can be regarded as the envelope constant. There is a simple illustration of how the algorithm works in Figure 5.1. Initially, sample a candidate $\mathbf{\Psi}^{(i)}$ from the proposal distribution $q(\mathbf{\Psi})$, and sample a variable u from a uniform distribution between 0 and 1. If $uMq(\mathbf{\Psi}^{(i)}) < \pi(\mathbf{\Psi}^{(i)})$, then the candidate is accepted, otherwise it is rejected. The accept/reject procedure [126,127] is described in Algorithm 5.1.

This simple method has some limitations, which are mentioned in [128,129]. First, M is not easy to be found to cover the whole region of $\pi(\mathbf{\Psi})$. Second, if M is too large, the acceptance probability in (5.8) will be too small. This will make the algorithm inefficient in high-dimensional situations.

$$Pr(\mathbf{\Psi}_{accept}) = Pr\left(u < \frac{\pi(\mathbf{\Psi})}{Mq(\mathbf{\Psi})}\right) = \frac{1}{M}. \tag{5.8}$$

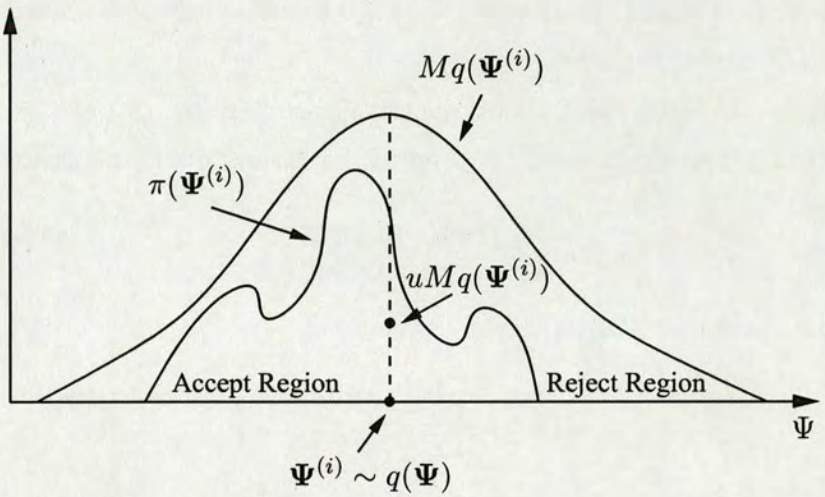


Figure 5.1: *Illustration of the rejection sampling algorithm.*

5.2.4 Importance sampling

In some scenarios, the envelope constant, M, of the distribution of interest is unknown or it cannot be estimated. An *importance sampling* approach needs to be adopted. As an alternative to the direct sampling, which has been introduced since 1940's in [130,131]. Denote an arbitrary importance proposal distribution by $q(\Psi)$. The integral, I_f, in (5.1), and a possible Monte Carlo estimate of it, $\hat{I}_f$, in (5.4) can be rewritten as:

$$I_f = \int f(\Psi)\omega(\Psi)q(\Psi)d\Psi \tag{5.9}$$

$$\hat{I}_f = \sum_{i=1}^{N} f(\Psi^{(i)})\omega(\Psi^{(i)}), \tag{5.10}$$

where $\omega(\Psi)$ is a set of *importance weights*, and $\omega(\Psi^{(i)})$ is the ith element in $\omega(\Psi)$. It is defined as:

$$\omega(\Psi^{(i)}) \triangleq \frac{\pi(\Psi^{(i)})}{q(\Psi^{(i)})}. \tag{5.11}$$

If the target distribution, $\pi(\Psi)$, is completely known, the weights, $\omega(\Psi)$, would be normalized already. Then the numerical approximation of $\pi(\Psi^{(i)})$ can be obtained with N i.i.d. samples:

$$\hat{\pi}(\Psi^{(i)}) = \sum_{i=1}^{N} \omega(\Psi^{(i)})\delta_{\Psi}(\Psi^{(i)}). \tag{5.12}$$

Unfortunately, $\hat{I}_f$ is biased. Only if $N \to +\infty$, the expectation of $\hat{I}_f$ asymptotically converges to I_f, denoted as: $\lim_{N \to +\infty} \hat{I}_f = I_f$.

Nevertheless, in many applications, the total probability of $\pi(\mathbf{\Psi})$ is impossible to evaluate, and $\omega(\mathbf{\Psi})$ is only known up to a normalizing constant. In this case, I_f can be rewritten as:

$$I_f = \frac{\int f(\mathbf{\Psi})\omega(\mathbf{\Psi})q(\mathbf{\Psi})d\mathbf{\Psi}}{\int \omega(\mathbf{\Psi})q(\mathbf{\Psi})d\mathbf{\Psi}}, \tag{5.13}$$

and the normalization of weights can be denoted as:

$$\tilde{\omega}(\mathbf{\Psi}^{(i)}) = \frac{\omega(\mathbf{\Psi}^{(i)})}{\sum_{i=1}^{N} \omega(\mathbf{\Psi}^{(i)})}. \tag{5.14}$$

Therefore, the approximation of the target distribution, $\pi(\mathbf{\Psi})$, becomes:

$$\tilde{\pi}(\mathbf{\Psi}^{(i)}) = \sum_{i=1}^{N} \tilde{\omega}(\mathbf{\Psi}^{(i)})\delta_{\Psi}(\mathbf{\Psi}^{(i)}). \tag{5.15}$$

With a finite number of samples from $\pi(\mathbf{\Psi})$ using the approximation in (5.15), the estimates of I_f, i.e. $\tilde{I}_f$, can be presented as:

$$\tilde{I}_f = \frac{\frac{1}{N}\sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)})\omega(\mathbf{\Psi}^{(i)})}{\frac{1}{N}\sum_{i=1}^{N} \omega(\mathbf{\Psi}^{(i)})} = \sum_{i=1}^{N} \tilde{\omega}(\mathbf{\Psi}^{(i)}) f(\mathbf{\Psi}^{(i)}). \tag{5.16}$$

$\tilde{I}_f$ is also asymptotically converges to I_f if $N \to +\infty$. Moreover, the estimator $\tilde{I}_f$ has been shown to have better performance than $\hat{I}_f$ in some situations under squared error loss [126].

There are other variations based on this *importance sampling* method. For example, *adaptive importance sampling* originates from the structural safety literature [132], and widely applied to the communications literature [133]. It incorporates additional parameter, ξ, into $q(\mathbf{\Psi}, \xi)$, and adapts ξ during the simulation. Another example is known as *sampling importance resampling (SIR)* [134]. It allows obtaining M i.i.d. samples from $\hat{\pi}(\mathbf{\Psi})$. Brief descriptions about these methods can be found in [100, 128].

Even with adaptation and other improved sampling methods, obtaining a proposal distribution that is easy to sample from and achieving a good approximation result is still a trade-off. As a consequence, more sophisticated sampling algorithms based on Markov chains are required. They play a more and more important role in sampling from a complicated distribution of interest.

5.3 Markov chain Monte Carlo algorithms

MCMC methodology is generally Monte Carlo numerical integration using Markov chains. It enables simulation from a distribution by regarding it as a limiting distribution of a Markov chain and simulating from the chain until it approaches equilibrium. To be more specific, MCMC is a strategy for generating samples $\mathbf{\Psi}$ while exploring the state space using a Markov chain mechanism. It is constructed so that the samples $\mathbf{\Psi}$ can mimic samples drawn from the target distribution $\pi(\mathbf{\Psi})$. Combined with the Bayesian approach, MCMC algorithms can integrate the posterior distribution over undesired model parameters and make inference about other desired parameters with the observational data [129]. In this way, they provide a solution to the difficult problem of simulation from a high-dimensional distribution of the unknown parameters in complex models. Investigations from basic concepts to advanced algorithms are presented in [99,100]. A detailed review of MCMC algorithms can also be found in [128].

The following subsections start with some background information about Markov chains in Section 5.3.1, followed by two popular algorithms, MH algorithm in Section 5.3.2 and Gibbs sampling in Section 5.3.3. If the dimension of parameters is not fixed, a rjMCMC algorithm is introduced in Section 5.3.5.

5.3.1 Basic concepts on Markov chains

In mathematics, a Markov chain is a stochastic process, which deals with characterisation of sequences of random variables and is described in terms of states in discrete time [100]. Denote a stochastic process $\mathbf{\Psi}$ on finite state spaces, which takes s discrete values, $\mathbb{S} = [\mathcal{S}_{(1)}, \mathcal{S}_{(2)}, \cdots, \mathcal{S}_{(s)}]$. When $\mathbf{\Psi}^{(i)} = \mathcal{S}_{(q)}$, the process is said to be in state q. If the probability of the next state $\mathbf{\Psi}^{(i+1)}$ depends only on the present state $\mathbf{\Psi}^{(i)}$ given all previous values of the process, as in (5.17), it is called a first order Markov chain.

$$P(\mathbf{\Psi}^{(i+1)}|\mathbf{\Psi}^{(1)}, \mathbf{\Psi}^{(2)}, \cdots, \mathbf{\Psi}^{(i)}) = P(\mathbf{\Psi}^{(i+1)}|\mathbf{\Psi}^{(i)}). \tag{5.17}$$

Moreover, denote the transition probability from state $\mathcal{S}_{(m)}$ to $\mathcal{S}_{(n)}$ as $P_{m,n}(\mathbf{\Psi}^{(i+1)}) = P(\mathbf{\Psi}^{(i+1)} = \mathcal{S}_{(n)}|\mathbf{\Psi}^{(i)} = \mathcal{S}_{(m)})$. Assume the transition probabilities are stationary over time. Then a complete set of the probabilities can be represented by a stochastic tran-

sition matrix $\mathbf{T}$ as follows:

$$\mathbf{T} = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,s} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ P_{s,1} & P_{s,2} & \cdots & P_{s,s} \end{bmatrix}. \tag{5.18}$$

If $\mathbf{T} \triangleq \mathbf{T}(\mathbf{\Psi}^{(i+1)}|\mathbf{\Psi}^{(i)})$ remains invariant for all i, the Markov chain is called a homogeneous chain. Based on the transition matrix $\mathbf{T}$, a general expression for the probability of $\mathbf{\Psi}$ being in state $\mathcal{S}_{(n)}$ from state $\mathcal{S}_{(m)}$ after l iterations, can be written as:

$$P_{m,n}^{(l)}(\mathbf{\Psi}^{(i)}) = P(\mathbf{\Psi}^{(i)} = \mathcal{S}_{(n)}|\mathbf{\Psi}^{(i-l)} = \mathcal{S}_{(m)}) = \mathbf{T}^l P(\mathbf{\Psi}^{(i-l)}|\mathcal{S}_{(m)}). \tag{5.19}$$

If $n \rightarrow +\infty$, the Markov chain will converge, with some certain conditions, to a limiting distribution, which is independent of the initial state. When the chain has converged, the states of the resulting chain are distributed according to the limiting distribution $\mathbf{T}\pi = \pi$. Therefore, if the limiting distribution of a Markov chain is the desired posterior distribution, $\pi(\mathbf{\Psi})$, after some states, which are known as the 'burn-in' period of the Markov chain, have been discarded, the rest of states of the chain become samples from this posterior distribution $\pi(\mathbf{\Psi})$.

In continuous state spaces, the transition matrix $\mathbf{T}$ becomes an integral kernel K_c and the invariant distribution $\pi(\mathbf{\Psi})$ becomes an eigenfunction [1] [128]:

$$\int \pi(\mathbf{\Psi}^{(i)}) K_c(\mathbf{\Psi}^{(i+1)}|\mathbf{\Psi}^{(i)}) d\mathbf{\Psi}^{(i)} = \pi(\mathbf{\Psi}^{(i+1)}), \tag{5.20}$$

where K_c is the conditional density of $\mathbf{\Psi}^{(i+1)}$ given the value $\mathbf{\Psi}^{(i)}$.

For any starting point, a chain will converge to the desired density function as long as $\mathbf{T}$ is a stochastic transition matrix that obeys the following properties [128, 129]:

- Irreductibility: Within finite iterations, the Markov chain can reach any state with positive probability from all starting points.

- Aperiodicity: The chain should not get trapped in cycles.

[1] An eigenfunction of a linear operator, $\mathbf{A}$, is defined as a non-zero function, f, in some function space that returns from the operator exactly as it is, i.e. $\mathbf{A}f = \lambda f$, for some scalar, λ, which is the corresponding eigenvalue.

- Invariance: All states in the Markov chain have reached a limiting distribution and are distributed according to $\mathbf{T}\pi = \pi$.
- Reversibility: The probability of a transition of the Markov chain from one state to another equals to the probability of a transition in the reverse direction. It is a sufficient, but not necessary condition for the states of the chain to be in a limiting distribution. It can be denoted as:

$$\pi(\mathbf{\Psi}^{(i)})\mathbf{T}(\mathbf{\Psi}^{(i-1)}|\mathbf{\Psi}^{(i)}) = \pi(\mathbf{\Psi}^{(i-1)})\mathbf{T}(\mathbf{\Psi}^{(i)}|\mathbf{\Psi}^{(i-1)}). \tag{5.21}$$

- Recurrence: All states can be reached infinitely often from all starting points.

All MCMC algorithms have been designed to satisfy these constraints [128].

5.3.2 Metropolis-Hasting algorithm

The name of the MH algorithm originates from papers by Metropolis in [135] in 1953 and Hastings in [136] in 1970. It is the most popular MCMC technique and most of the practical MCMC algorithms can be interpreted as special cases or extensions of this algorithm [128].

5.3.2.1 Procedure and properties of the MH algorithm

Denote a candidate function, i.e. the proposal distribution, by $q(\cdot)$, which is similar to the importance function used in importance sampling. $\pi(\cdot)$ indicates the target distribution as usual. Assume a Markov chain is in state $\mathbf{\Psi}$. The candidate $\mathbf{\Psi}^*$ for the next state can be obtained by sampling from the proposal conditional distribution $q(\mathbf{\Psi}^*|\mathbf{\Psi})$. Then the candidate will be accepted with a probability defined as:

$$\alpha(\mathbf{\Psi}^*, \mathbf{\Psi}) = \min\{1, r(\mathbf{\Psi}^*, \mathbf{\Psi})\}, \tag{5.22}$$

where $r(\mathbf{\Psi}^*, \mathbf{\Psi})$ is the acceptance ratio defined as:

$$r(\mathbf{\Psi}^*, \mathbf{\Psi}) = \frac{\pi(\mathbf{\Psi}^*)q(\mathbf{\Psi}|\mathbf{\Psi}^*)}{\pi(\mathbf{\Psi})q(\mathbf{\Psi}^*|\mathbf{\Psi})}. \tag{5.23}$$

If the candidate is accepted, the Markov chain then moves towards the new state $\mathbf{\Psi}^*$; otherwise it remains at the current state $\mathbf{\Psi}$. The procedure can be summarized in Algorithm 5.2.

Algorithm 5.2: Metropolis-Hasting Algorithm

Initialize $\boldsymbol{\Psi}^{(0)}$.
for $i \geq 0$ **do**
  Sample $u \sim U_{[0,1]}$;
  Sample $\boldsymbol{\Psi}^* \sim q(\boldsymbol{\Psi}^*|\boldsymbol{\Psi}^{(i)})$;
  Evaluate the acceptance probability

$$\alpha(\boldsymbol{\Psi}^*, \boldsymbol{\Psi}^{(i)}) = \min\left\{1, \frac{\pi(\boldsymbol{\Psi}^*)q(\boldsymbol{\Psi}^{(i)}|\boldsymbol{\Psi}^*)}{\pi(\boldsymbol{\Psi}^{(i)})q(\boldsymbol{\Psi}^*|\boldsymbol{\Psi}^{(i)})}\right\}. \tag{5.24}$$

  if $u \leq \alpha(\boldsymbol{\Psi}^*, \boldsymbol{\Psi}^{(i)})$ **then**
    The chain takes the new state $\boldsymbol{\Psi}^{(i+1)} = \boldsymbol{\Psi}^*$;
  else
    It remains at the current state $\boldsymbol{\Psi}^{(i+1)} = \boldsymbol{\Psi}^{(i)}$;
  end
end

The transition kernel for the MH algorithm is created:

$$\mathcal{K}(\boldsymbol{\Psi}^{(i+1)}|\boldsymbol{\Psi}^{(i)}) = q(\boldsymbol{\Psi}^{(i+1)}|\boldsymbol{\Psi}^{(i)})\alpha(\boldsymbol{\Psi}^{(i)}, \boldsymbol{\Psi}^{(i+1)}) + \delta_{\boldsymbol{\Psi}^{(i)}}(\boldsymbol{\Psi}^{(i+1)})r(\boldsymbol{\Psi}^{(i)}), \tag{5.25}$$

where $\delta(\cdot)$ is the Dirac delta function and $r(\boldsymbol{\Psi}^{(i)})$ is related to the rejection:

$$r(\boldsymbol{\Psi}^{(i)}) = \int q(\boldsymbol{\Psi}^*|\boldsymbol{\Psi}^{(i)})(1 - \alpha(\boldsymbol{\Psi}^*, \boldsymbol{\Psi}^{(i)}))d\boldsymbol{\Psi}^*. \tag{5.26}$$

The constructed transition kernel satisfy the property of reversibility, i.e.:

$$\pi(\boldsymbol{\Psi}^{(i)})\mathcal{K}(\boldsymbol{\Psi}^{(i+1)}|\boldsymbol{\Psi}^{(i)}) = \pi(\boldsymbol{\Psi}^{(i+1)})\mathcal{K}(\boldsymbol{\Psi}^{(i)}|\boldsymbol{\Psi}^{(i+1)}), \tag{5.27}$$

and thus the MH algorithm regards $\pi(\boldsymbol{\Psi})$ as the invariant distribution. The proof of this detailed balance condition can be found in [122]. Furthermore, the MH algorithm requires $q(\boldsymbol{\Psi})$ be strictly positive on the support of $\pi(\boldsymbol{\Psi})$, and always allows for rejection, so as to satisfy the irreductibility and the aperiodicity properties of the Markov chain [129].

One major advantage of the MH algorithm is that the normalizing constant of the target distribution is not required. As the target distribution is present in ratios, the unknown normalizing constant will cancel out. Furthermore, one significant issue in the MH algorithm lies in the choice of the proposal distribution $q(\cdot)$. In [128], there is a

simple illustration showing that the choice of the proposal standard deviation σ_p of the proposal distribution will have influence on the property of the chain. It is pointed out when the target distribution has multiple modes, if σ_p is too narrow, only one mode of the target distribution might be visited; if σ_p is too wide, the rejection rate will be very high. Only when all the modes are visited and the acceptance probability is high, the chain is said to be mixed well. In order to have a well-mixed chain, $q(\cdot)$ should allow the chain to explore the entire state space and with high acceptance probability.

Metropolis algorithm is a special case of the MH algorithm. It assumes the proposal distribution $q(\cdot)$ has a symmetric property, i.e. $q(\mathbf{\Psi}^*|\mathbf{\Psi}^{(i)}) = q(\mathbf{\Psi}^{(i)}|\mathbf{\Psi}^*)$. In this case, the acceptance probability reduces to

$$\alpha(\mathbf{\Psi}^*, \mathbf{\Psi}^{(i)}) = \min\left\{1, \frac{\pi(\mathbf{\Psi}^*)}{\pi(\mathbf{\Psi}^{(i)})}\right\}, \tag{5.28}$$

and it is independent of $q(\cdot)$.

5.3.2.2 Hybrid kernels in the MH algorithm

A powerful property of MCMC algorithms is that it is possible to combine several samplers into mixtures and cycles of the individual samplers [128, 137].

In the case of mixture transition kernel with component transition kernels $\mathcal{K}_i, i = 1, \ldots, r$, ($r$ is the number of components in the mixture), the new hybrid kernel $\mathcal{K}_m$ can be obtained by taking:

$$\mathcal{K}_m = \sum_{i=1}^{r} \omega_i \mathcal{K}_i, \tag{5.29}$$

where ω_i, $(i = 1, \ldots, r)$ are the weights and satisfy $\omega_i \geq 0$ and $\sum_{i=1}^{r} \omega_i = 1$. The property of the mixture transition kernel, $\mathcal{K}_m$, has the same properties of the transition kernels $\mathcal{K}_i$. As the component transition kernels $\mathcal{K}_i$ have a common equilibrium distribution $\pi(\cdot)$, the mixture kernel also defines the transition kernel of a Markov chain with equilibrium distribution $\pi(\cdot)$. In addition, if one of the component transition kernels $\mathcal{K}_i$ is irreducible and aperiodic, then the mixture kernel $\mathcal{K}_m$ is irreducible and aperiodic.

With regard to the cycle transition kernel $\mathcal{K}_c$, it also has component transition kernels $\mathcal{K}_i$, $(i = 1, \ldots, r)$. Each of them has the same equilibrium distribution $\pi(\cdot)$. In an iteration, the new chain is performed after all moves, guided by the component

kernels, are gone through. Thus the cycle hybrid kernel can be obtained by taking:

$$\mathcal{K}_c = \prod_{i=1}^{r} \mathcal{K}_i. \tag{5.30}$$

Similar to the mixture kernel, many properties of the component transition kernels $\mathcal{K}_i$ are passed on to the cycle hybrid kernel $\mathcal{K}_c$, such as the common equilibrium distribution $\pi(\cdot)$. However, unlike the mixture kernels, only if all the component kernels are irreducible and aperiodic, then the cycle kernel is irreducible and aperiodic.

As an application of the hybrid kernels of the MH algorithm, the spectral estimation using a parametric model within Bayesian framework utilizes a mixture of kernel with two component transition kernels, i.e. global and local proposal kernels. It has been mentioned in [138, 139] that mixture of kernels can incorporate the global proposal to explore most regions of the state space and the local proposal to discover finer details. It is particularly useful when the target distribution has many narrow peaks. The global proposal locks into the peaks while the local proposal allows to explore the state space around each peak. In [33], the frequency spectrum based on the Fourier transform is adopted as a global proposal, $\mathcal{K}_1$, and a random walk algorithm is used as a local proposal, $\mathcal{K}_2$, which has been demonstrated to be a high precision frequency estimator. The mixture hybrid kernel can be written as $\lambda\mathcal{K}_1 + (1-\lambda)\mathcal{K}_2$, for $0 \leq \lambda \leq 1$. They have a common target distribution, which is the posterior distribution of the spectral content.

Cycles of kernels can split a multivariate state vector into components or blocks that can be updated separately [128]. The samplers will mix more quickly by blocking highly correlated variables. A popular cycle of the MH kernels is Gibbs sampler, which becomes more and more popular in many applications, and will be introduced later in Section 5.3.3.

5.3.3 Gibbs sampling

Gibbs sampler can be regarded as another special case of the MH algorithm where the transition kernel is constructed by the full conditional distributions. [2] It allows to break down the problem by drawing samples from a multivariate density into drawing samples successively from densities of smaller dimensionality. The major feature is

[2]The full conditional distribution of a parameter is defined as the distribution of a parameter conditioned on all other parameters.

that it does not require the selection of a proposal distribution and everything is fixed by the target distribution $\pi(\cdot)$ [140, 141].

5.3.3.1 Procedure of the algorithm

Define a parameter vector $\mathbf{\Psi}$ with a length of N and the full conditionals:

$$\pi(\Psi_j^{(i)}|\mathbf{\Psi}_{-j}^{(i)}) \triangleq \pi(\Psi_j^{(i)}|\Psi_1^{(i)}, \cdots, \Psi_{j-1}^{(i)}, \Psi_{j+1}^{(i)}, \cdots, \Psi_N^{(i)}), \tag{5.31}$$

where $j = 1, \ldots, N$. Rather than sampling from a complex N dimensional distribution, the problem is simplified to sampling for N times from one dimensional conditional distribution. In Gibbs sampling, the proposal distributions for $j = 1, \ldots, N$ are specified as:

$$q(\mathbf{\Psi}_j^*|\mathbf{\Psi}_j^{(i)}) = \begin{cases} \pi(\Psi_j^*|\mathbf{\Psi}_{-j}^{(i)}) & \text{If } \mathbf{\Psi}_{-j}^* = \mathbf{\Psi}_{-j}^{(i)} \\ 0 & \text{Otherwise.} \end{cases} \tag{5.32}$$

By substituting (5.32) into (5.24), the corresponding acceptance probability can be simplified as:

$$\begin{aligned} \alpha(\mathbf{\Psi}_j^*, \mathbf{\Psi}_j^{(i)}) &= \min\left\{1, \frac{\pi(\mathbf{\Psi}_j^*)q(\mathbf{\Psi}_j^{(i)}|\mathbf{\Psi}_j^*)}{\pi(\mathbf{\Psi}_j^{(i)})q(\mathbf{\Psi}_j^*|\mathbf{\Psi}_j^{(i)})}\right\} \\ &= \min\left\{1, \frac{\pi(\mathbf{\Psi}_j^*)\pi(\Psi_j^{(i)}|\mathbf{\Psi}_{-j}^{(i)})}{\pi(\mathbf{\Psi}^{(i)})\pi(\Psi_j^*|\mathbf{\Psi}_{-j}^*)}\right\} \\ &= \min\left\{1, \frac{\pi(\Psi_j^*|\mathbf{\Psi}_{-j}^{(i)})\pi(\Psi_j^{(i)}|\mathbf{\Psi}_{-j}^{(i)})}{\pi(\Psi_j^{(i)}|\mathbf{\Psi}_{-j}^{(i)})\pi(\Psi_j^*|\mathbf{\Psi}_{-j}^*)}\right\} \\ &= 1. \end{aligned} \tag{5.33}$$

Therefore, Gibbs sampling can be regarded as an MH algorithm with the acceptance probability of 1. If all the full conditional posterior distributions are available and belong to the family of standard distributions, Gibbs sampling is able to draw new samples directly. However, if the analytical form of the full conditional posterior distributions is not available, samples can be drawn with MH steps embedded within the Gibbs algorithm [128]. The procedure of Gibbs sampling algorithm is described in Algorithm 5.3:

Algorithm 5.3: Gibbs sampling Algorithm

Initialize $\mathbf{\Psi}^{(0)}$;
for $i \geq 0$ **do**
  Sample $\Psi_1^{(i+1)} \sim \pi(\Psi_1|\Psi_2^{(i)}, \Psi_3^{(i)}, \ldots, \Psi_N^{(i)})$;
  Sample $\Psi_2^{(i+1)} \sim \pi(\Psi_2|\Psi_1^{(i+1)}, \Psi_3^{(i)}, \ldots, \Psi_N^{(i)})$;
  $\vdots$
  Sample $\Psi_j^{(i+1)} \sim \pi(\Psi_j|\Psi_1^{(i+1)}, \ldots, \Psi_{j-1}^{(i+1)}, \Psi_{j+1}^{(i)}, \ldots, \Psi_N^{(i)})$;
  $\vdots$
  Sample $\Psi_N^{(i+1)} \sim \pi(\Psi_N|\Psi_1^{(i+1)}, \Psi_2^{(i+1)}, \ldots, \Psi_{N-1}^{(i+1)})$;
end

5.3.3.2 Application in spectral analysis

Dou and Hodgson applied Gibbs sampling algorithms in spectral analysis in [142]. They examined a signal with multiple frequencies. The posterior distribution for the frequencies are obtained using Bayesian inference. Although they are complicated, they are in the family of standard distributions and thus Gibbs sampling can be used. Moreover, Dou and Hodgson extended the signal model from harmonic models to more complicated models in [143], such as periodic but non-harmonic signals, signals with decays and signals with chirp. It was demonstrated that Gibbs sampling can also be applied to these models to achieve the estimated frequencies in the signals.

Unfortunately, in most of the practical applications, it is not always known which model applies. Particularly, the number of frequency components in a signal, indicating the model order of the signal, is usually unknown. Dou and Hodgson developed an intuitive way to tackle the problem in [143]. They assumed a model space with all possible models and then tried to find the best choice. In particular, a uniform prior probability to the observation on the model space is first assigned. Then the posterior probabilities of the models in the model space are compared by calculating the relative posterior probabilities, i.e. the posterior odds ratio. The examples have demonstrated that the method works fine but has large a computational complexity. Also, a large amount of memory is needed to store the probabilities. Apart from the above mentioned method used in determining the model order of harmonic signals, a better technique, i.e. a rjMCMC algorithm, can be introduced to jointly estimate the model order and the frequency components in the signal. The principle of the rjMCMC algorithm will be discussed in Section 5.3.5.

5.3.4 Convergence diagnostics

In theory, a sample from the desired distribution of interest $\pi(\cdot)$ can only be obtained when the number of iterations of a Markov chain approaches infinity. In practice, it is impossible and a sample obtained at a sufficiently large iteration is taken. Unfortunately, it is hard to determine how large this iteration should be, which leads to a convergence problem.

There are two major ways to deal with the convergence problem. One is more theoretical, and tries to measure the total variation distance between the distribution of a chain at each iteration and its limiting distribution. This approach attracts many studies in [144–146]. However, it is pointed out in [147] that, the results are not satisfactory in practical work. The other one is more empirical and practical. It analyzes the property of the observed output from a chain to check whether the chain converges or not from a statistical perspective. Nevertheless, it is only based on observations from the chain and thus can never guarantee the convergence. There are comparative and illustrative reviews about these methods in [147,148]. Although two approaches complement each other, the theoretical results are more difficult to attain and thereby, the convergence diagnostics according to the statistical properties of the observed chain will be adopted in this thesis.

By exploring the statistical properties of the observed Markov chain, the convergence can be checked by a few informal monitors or some formal convergence methods. The informal checks of the convergence are usually simple and easy to deal with. For example, one informal check is based on graphical representation techniques, which are suggested in [149]. With a single chain, a trajectory of the chain exhibiting the same qualitative behaviour through iterations, or an asymptotic behaviour over many successive iterations indicates convergence. Moreover, the visual impression of the convergence can be reinforced when the chain started at various values oscillate in the same region or concentrate around the same value. In opposition to the informal checks, the formal convergence methods attempt to decide whether the convergence can be assumed to hold rather than determining the run length to achieve the convergence. Most of these methods are described in the review papers in [147,150].

Furthermore, there are many suggestions for improving the convergence of a chain to the equilibrium distribution in the literature. They are divided into two groups: alterations in the chain and alterations in the equilibrium distribution. The former tech-

niques include blocking parameters [151, 152] and reparametrization [100, 153]. The latter techniques have a famous application, which is obtained based on the optimization technique known as simulated annealing [154]. Moreover, additional variables can be introduced as auxiliary variables [155] in order to remove the correlation in the chain and thus hasten the convergence.

These approaches for convergence diagnostics discussed above can be applied to all varieties of MCMC algorithms. Nevertheless, the investigations of Markov chain convergence is another hot research topic in developing MCMC algorithms. It will not be examined in this thesis, and the informal checks are adopted to monitor the possibility and the velocity of the chain convergence.

5.3.5 Reversible jump MCMC

In previous sections, the sampling techniques all assume the model order dimension is known or fixed. When considering building MCMC algorithms into model order detection, the main difficulty lies in that the sampler is able to jump from one subspace to another while keeping the correct target distribution. Some conventional approaches which are used in computer vision problems, are based on the jump-diffusion samplers [156]. However, it only approximately maintains the reversibility property of the constructed Markov chain. Moreover, the range of transitions is limited and restricted for general Bayesian computation. Another solution, which is based on sampling from the product space, has been described in [157]. In this approach, the appropriate model order can be detected by being sampled from the subspace corresponding to a range of model orders independently. Nevertheless, the update of irrelevant parameters in some models that have a low posterior probability will make this approach inefficient and computationally intensive.

Green [122] developed a general framework, known as the reversible jump MCMC algorithm in 1995. It allows the sampling process to jump between subspaces of different dimensions and preserve the reversibility at the same time. Moreover, it allows for the joint estimation of the model order and the sampling of a parameter from its posterior distribution. Furthermore, it avoids the expensive search over the full product parameter space. Denote a discrete variable k as the model order, which is one of the parameters of interest. The whole parameter space can be represented by $\bigcup_{k=0}^{k_{max}} k \times \mathbf{\Phi}_k$, in which $\mathbf{\Phi}_k$ is a set of parameters of the model with order k, and k_{max} is the maximum allowable model order. Like in the MH algorithm, denote a proposal distribution by

$q(\cdot)$ and a candidate vector $\mathbf{\Phi}_k^*$. In each iteration, the candidate $\mathbf{\Phi}_k^*$ is accepted with a probability in the form of:

$$\alpha_{accept} = \min\{1, (\text{likelihood ratio}) \times (\text{prior ratio}) \times (\text{proposal ratio}) \times (\text{Jacobian})\}. \tag{5.34}$$

This expression can be derived from the acceptance probability for the proposed transition or switching between two subspaces, which can be found in [122]. In (5.34), the product of likelihood ratio and prior ratio equals to the corresponding posterior ratio, i.e. the ratio of $\pi(\cdot)$ with different model order. Therefore, it can be re-written as:

$$\alpha\left((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k)\right) = \min\left\{1, r\left((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k)\right)\right\}, \tag{5.35}$$

where the acceptance ratio is defined as:

$$r((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k)) = \frac{\pi(k^*, \mathbf{\Phi}_k^*) q(k, \mathbf{\Phi}_k)}{\pi(k, \mathbf{\Phi}_k) q(k^*, \mathbf{\Phi}_k^*)} \times \mathbf{J}((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k)). \tag{5.36}$$

In order to ensure the reversibility condition, the Jacobian term $\mathbf{J}((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k))$ in (5.36) is required to calculate the transformation probability between spaces of different dimensions. According to [158], it can be given by :

$$\mathbf{J}((k^*, \mathbf{\Phi}_k^*), (k, \mathbf{\Phi}_k)) = \left| \frac{\partial(k^*, \mathbf{\Phi}^*)}{\partial(k, \mathbf{\Phi})} \right|, \tag{5.37}$$

where $|\cdot|$ denotes the determinant operation.

Apart from the requirements for calculating the acceptance probability, the proposed candidate function also needs to be selected appropriately. There is flexibility in selecting various moves/functions. In generic reversible jump MCMC technique, the most widely used functions include *merge* and *split* moves, *birth* and *death* moves, etc. Specifically, a *merge* move is defined as combining two nearby components and a *split* move is defined as breaking a component into two nearby ones. Moreover, *birth* and *death* moves are carried out by birth of a component and death of a component. They are both complementary moves that ensure the reversibility. A lot of investigations and discussions about different candidate moves can be found in [33, 158, 159].

Nevertheless, in order to simplify the dimension-changing problem in application of the reversible jump technique, the most popular and commonly used moves are *birth*, *death* and *update* moves in the spectral analysis. During the *birth* move, the algo-

Algorithm 5.4: Reversible jump MCMC Algorithm

```
Initialize (k^(0), Φ_k^(0)) ;
for i = 1 to numIteration do
    Sample u ~ U_[0,1] ;
    Calculate the probabilities b_k and d_k according to (5.38) and (5.39) ;
    if u ≤ b_k(i) then
        Execute a birth move ;
    else if u ≤ b_k(i) + d_k(i) then
        Execute a death move ;
    else
        Execute an update move ;
    end
end
```

rithm proposes a candidate of higher dimension whereas in the *death* move, it proposes a candidate of lower dimension. Bear these basic definitions and concepts in mind, the state transitions can be described in more detail. To be more specific, there are three available move types: $\mathbf{\Phi}_k^*$, $\mathbf{\Phi}_{k+1}$ and $\mathbf{\Phi}_{k-1}$, indicating the *update* move, *birth* move and *death* move respectively, compared to the current state $\mathbf{\Phi}_k$. For each transition, an independent random choice is made. The probabilities for choosing each single move can be denoted as u_k, b_k and d_k, which will satisfy $u_k + b_k + d_k = 1$ for all k. In accordance with [122], the probabilities of *birth* and *death* moves can be defined as:

$$b_k = \begin{cases} c \min\left\{1, \frac{p(k+1)}{p(k)}\right\} & 0 \leq k < k_{max} \\ 0 & k = k_{max} \end{cases} \tag{5.38}$$

$$d_k = \begin{cases} 0 & k = 0 \\ c \min\left\{1, \frac{p(k)}{p(k+1)}\right\} & 1 \leq k \leq k_{max} \end{cases}, \tag{5.39}$$

where $p(k)$ represents the prior distribution of the model parameter k. The constant c is a tuning factor which determines the ratio of the *update* move to the jump moves. $c = 0.5$ is chosen so that the probability of a jump move is between 0.5 and 1 at every iteration. This choice ensures $b_k p(k) = d_{k+1} p(k+1)$, which could guarantee certain acceptance in the corresponding Hastings sampler [122]. In addition, the Jacobian term during the *birth/death* moves can be simplified to unity according to (5.37) [159]. The overall description of the rjMCMC algorithm is summarized in Algorithm 5.4.

Algorithm 5.5: Birth Move

1 Propose a new element Φ_e to form a new candidate vector

$$\mathbf{\Phi}_{k+1}^{(i+1)} = [\mathbf{\Phi}_k^{(i)}, \Phi_e]. \tag{5.44}$$

2 Evaluate α_{birth} according to (5.42) and (5.43) ;
3 Sample $u \sim U_{[0,1]}$;
4 **if** $u \leq \alpha_{birth}$ **then**
5 | The state of the Markov chain becomes $(k+1, \mathbf{\Phi}_{k+1}^{(i+1)})$;
6 **else**
7 | The state remains at $(k, \mathbf{\Phi}_k^{(i)})$;
8 **end**

5.3.5.1 Birth move

In the process of a *birth* move, the current state is denoted as $(k, \mathbf{\Phi}_k)$. The objective is to determine whether the next state is in $(k+1, \mathbf{\Phi}_{k+1})$ with a newly proposed element Φ_e. Φ_e is selected from a proposal distribution, $q(\cdot)$, and is inserted to the existing parameter vector to create a new one:

$$\mathbf{\Phi}_{k+1} = [\mathbf{\Phi}_k, \Phi_e]. \tag{5.40}$$

The proposal distribution conditioned on the current state $q(k+1, \mathbf{\Phi}_{k+1}|k, \mathbf{\Phi}_k)$ for the *birth* move can thus be written as:

$$q(k+1, \mathbf{\Phi}_{k+1}|k, \mathbf{\Phi}_k) = p(k+1)p(\Phi_e), \tag{5.41}$$

where $p(k+1)$ is the prior distribution of model parameter $k+1$ and $p(\Phi_e)$ is the distribution of the individual candidate parameter. Then the candidate state $(k+1, \mathbf{\Phi}_{k+1})$ is accepted with a probability:

$$\alpha_{birth} = \min\{1, r_{birth}\}, \tag{5.42}$$

where the acceptance ratio r_{birth} can be specified in accordance with (5.36):

$$r_{birth}((k+1, \mathbf{\Phi}_{k+1}), (k, \mathbf{\Phi}_k)) = \frac{\pi(k+1, \mathbf{\Phi}_{k+1})q(k, \mathbf{\Phi}_k)}{\pi(k, \mathbf{\Phi}_k)q(k+1, \mathbf{\Phi}_{k+1})}. \tag{5.43}$$

Therefore, the procedure of the *birth* move can be described in Algorithm 5.5.

Algorithm 5.6: Death Move

1 Randomly choose the j^{th} element $\Phi_j^{(i)}$ at the ith iteration to create a new candidate vector

$$\mathbf{\Phi}_k^{(i+1)} = [\mathbf{\Phi}_{1:j-1}^{(i)}, \mathbf{\Phi}_{j+1:k+1}^{(i)}]. \tag{5.48}$$

2 Evaluate α_{death} according to (5.43), (5.45) and (5.46) ;
3 Sample $u \sim U_{[0,1]}$;
4 **if** $u \leq \alpha_{death}$ **then**
5 | The state of the Markov chain becomes $(k, \mathbf{\Phi}_k^{(i+1)})$;
6 **else**
7 | The state remains at $(k+1, \mathbf{\Phi}_{k+1}^{(i)})$;
8 **end**

5.3.5.2 Death move

The reversibility property of a Markov chain with respect to moves across different subspaces is of significant importance in maintaining the invariant target distribution of the rjMCMC algorithm for the model order. A sufficient condition for reversibility with regard to model order is that the acceptance ratio for a *death* move needs to be defined as [122]:

$$r_{death} = \frac{1}{r_{birth}}. \tag{5.45}$$

In this way, the probability of moving from model order k to $k+1$ equals to that of moving from $k+1$ to k. In a *death* move, a parameter, $\Phi_j^{(i)}$, in the current state $(k+1, \mathbf{\Phi}_{k+1}^{(i)})$ is randomly chosen to be removed. With a certain acceptance probability:

$$\alpha_{death} = \min\{1, r_{death}\}, \tag{5.46}$$

the new state will become $(k, \mathbf{\Phi}_k^{(i+1)})$ at the next iteration. It is assumed that the removal of an element from the parameter vector is selected among the existing $k+1$ elements at random, thus the proposal distribution $q(k, \mathbf{\Phi}_k | k+1, \mathbf{\Phi}_{k+1})$ can be represented by:

$$q(k, \mathbf{\Phi}_k | k+1, \mathbf{\Phi}_{k+1}) = p(k) \times \frac{1}{k+1}, \tag{5.47}$$

where the definition of $p(k)$ is the same as the one used in (5.41). Similarly, the procedure of the *death* move can be described in Algorithm 5.6.

Algorithm 5.7: Update Move

1 Sample a candidate Φ_j^* from the proposal distribution $q(\mathbf{\Phi}^*|\mathbf{\Phi})$ to create a new candidate vector $\mathbf{\Phi}_k^*$ according to (5.49) ;
2 Evaluate the acceptance probability α_{update} according to (5.50) and (5.51);
3 Sample $u \sim U_{[0,1]}$;
4 **if** $u \leq \alpha_{update}$ **then**
5 | The state of the Markov chain becomes $(k, \mathbf{\Phi}_k^*)$;
6 **else**
7 | The state remains at $(k, \mathbf{\Phi}_k)$;
8 **end**

5.3.5.3 Update move

If neither a *birth* move nor a *death* move is chosen, an *update* move is performed. In this move, the dimension does not change thus the Jacobian term in (5.36) can be omitted. For a fixed k, a new candidate Φ_j^* can be drawn from the proposal distribution $q(\mathbf{\Phi}_k^*|\mathbf{\Phi}_k)$ conditioned on the current value $\mathbf{\Phi}_k$. The state changes from $(k, \mathbf{\Phi}_k)$ to $(k, \mathbf{\Phi}_k^*)$ only with the update of parameter vector $\mathbf{\Phi}_k^*$, which is given by:

$$\mathbf{\Phi}_k^* = [\mathbf{\Phi}_{1:j-1}^{(i)}, \Phi_j^*, \mathbf{\Phi}_{j+1:k}^{(i)}]. \tag{5.49}$$

The acceptance probability for an *update* move is stated as:

$$\alpha_{update} = \min\{1, r_{update}\}, \tag{5.50}$$

where the acceptance ratio is defined according to (5.36):

$$r_{update} = \frac{\pi(\mathbf{\Phi}_k^*)q(\mathbf{\Phi}_k|\mathbf{\Phi}_k^*)}{\pi(\mathbf{\Phi}_k)q(\mathbf{\Phi}_k^*|\mathbf{\Phi}_k)}. \tag{5.51}$$

The procedure of performing the *update* move is described in Algorithm 5.7.

5.4 Chapter summary

This chapter has presented detailed descriptions of a variety of approximation approaches for Bayesian inference and has provided a brief review of MCMC algorithms as sampling strategies from sophisticated distributions. The advantages and disadvantages of various techniques examined in this chapter provide a fundamental basis for sampling from the complicated posterior distributions in Chapter 6 and Chapter 7.

Therefore, further analysis for measured ultrasound echo signals can be made possible.

Chapter 6
Temporal and spectral estimation using a rjMCMC algorithm

A signal model was proposed for measured ultrasound echo signals and Bayesian inference was used to estimate the model parameters in Chapter 4. However, the posterior distribution for model parameters cannot be obtained analytically. Chapter 5 provided a general methodology to approximate the Bayesian computation. While in this chapter, the estimation of temporal and spectral contents of measured ultrasound echo signals using a rjMCMC algorithm is described. Moreover, by taking the estimated results from the *non-parametric* methods as initial values for the *parametric* method, a new estimation system is created. It is first evaluated using a simulated signal, and then applied to experimentally measured ultrasound echo signals to extract their characteristics in both the time and the frequency domains simultaneously.

6.1 Introduction

According to the general discussion of *parametric* estimation method in Chapter 4, an appropriate signal model for measured ultrasound echo signals is first selected, then the model parameters are estimated using Bayesian inference. As the obtained posterior distribution of parameters is not available in an analytical form, approximation methods will be employed to solve this problem. The whole estimation procedure can be divided into two steps. First of all, an estimation of pulse locations in a multiple pulse signal is carried out. Then an estimation of spectral content in each single pulse segment of the signal is performed. The two steps are implemented in a Bayesian framework by constructing a Markov chain and estimating the temporal and spectral contents in an iterative way.

During the process of iteration steps, the initial values for model parameters must be chosen properly, which is one of the most important factors in the estimation. In general, the simplest way is to choose the initial values randomly. However, it sometimes requires long iterations to converge to the true values of model parameters,

which may result in an intensive computation. If this happens, a fast convergence is necessary. In Bayesian data analysis, it is always useful to obtain a rough estimate of the location of the target distribution, i.e. a point estimate of the parameters in the model, first by using some simple and non-iterative technique before developing more elaborate approximations or complicated methods [99]. As a consequence, results from the *non-parametric* estimation obtained in Chapter 3 can be taken as the rough estimates of the parameters in the first place. Then these coarse estimates are updated iteratively using the constructed Markov chain. Finally the estimates will converge to the true values of the model parameters after a certain number of iterations. A proposed estimation system is displayed in Figure 6.1. It contains the *non-parametric* estimation part, which has been discussed in Chapter 3; and the *parametric* estimation part, which is based on the proposed sophisticated signal model in Chapter 4. This chapter will focus on estimating model parameters by approximation of Bayesian inference using a rjMCMC algorithm.

In this chapter, the spectral estimation in a single pulse segment and the temporal estimation in a full length signal model with multiple pulses will be examined in Section 6.2 and Section 6.3 respectively. Since the estimation system is created, the performance evaluation will be carried out in Sections 6.5. Finally, the estimation results of experimentally measured ultrasound echo signals will be presented in Section 6.6.

6.2 Spectral estimation in a single pulse segment

In Section 4.5, the full length signal model and the model for one single pulse in the signal have been both expressed in a vector-matrix form in (4.17a) on page 91 and (4.19) on page 92 respectively. This section will concentrate on a single pulse segment of a signal model using Bayesian inference to estimate its spectral content. Moreover, rjMCMC algorithms will be introduced to approximate the inference for parameters in a single pulse segment. A brief explanation of deriving the joint posterior distribution using Bayesian inference will be given in Section 6.2.1 and Section 6.2.2. Then they are followed by a detailed description of rjMCMC diversity steps in approximating the obtained posterior distribution, which will be studied in Section 6.2.3.

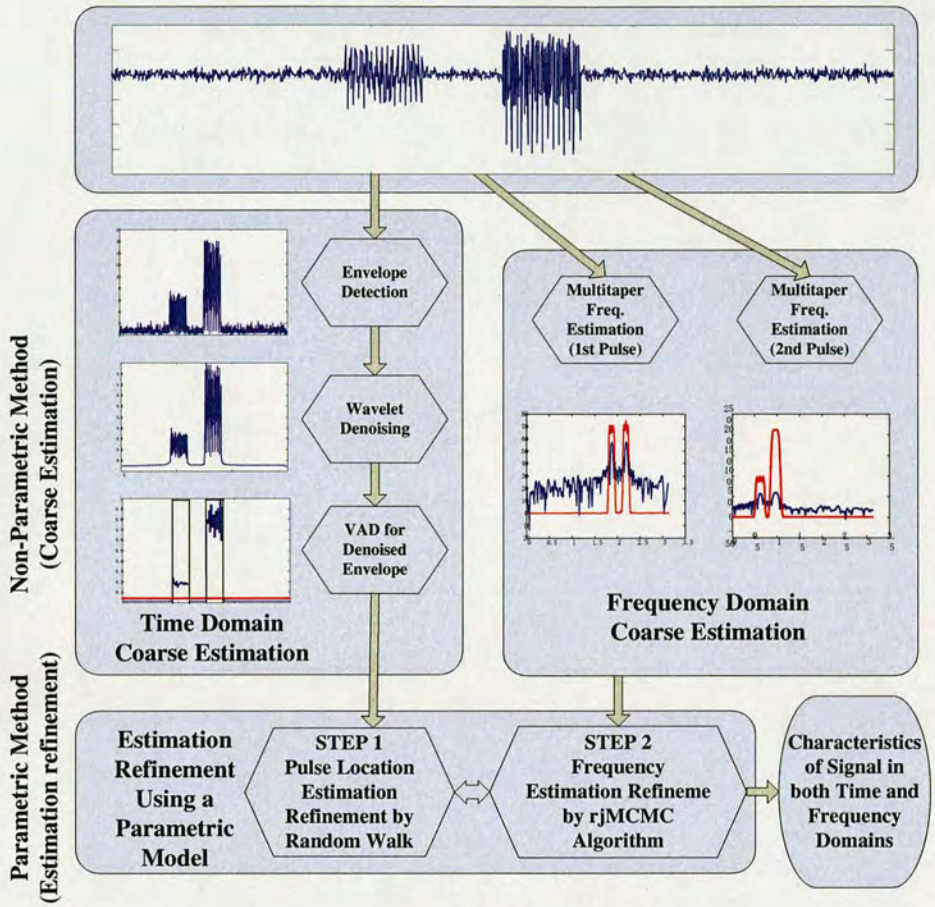


Figure 6.1: *Diagram for the procedure of the estimation system.*

6.2.1 Joint prior distribution for parameters in a single pulse

In Bayesian data analysis, the parameters of a signal model can be estimated based on their corresponding posterior distributions, which are computed with the given appropriate prior distributions [99]. These priors reflect the degree of belief of the relevant values of the parameters.

Conventionally, DAG [34,160,161] is a popular tool to illustrate a hierarchical structure of all parameters in a specific signal model. Figure 6.2 shows an DAG for the proposed model for a single pulse segment expressed in (4.19) on page 92. In this graph, random variables are represented by circular boxes, the fixed parameters are represented by square boxes, and the observed data is denoted by a rectangular box. To be more specific, v_i and γ_i, which are held fixed, are hyperparameters of the variable $\sigma_{k_i}^2$. Moreover, Λ_i is the hyperparameter of the number of sinusoids k_i, and δ_i^2 is the

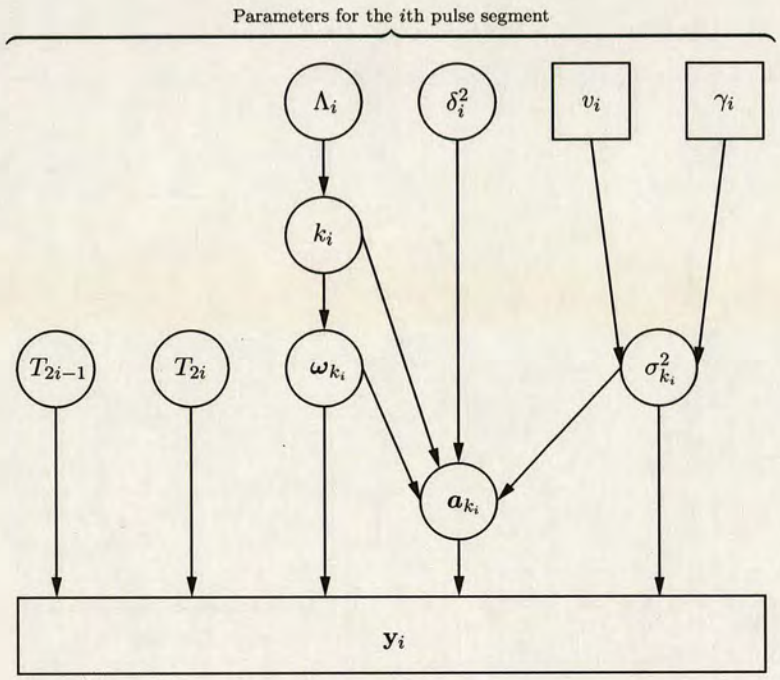


Figure 6.2: *DAG of hierarchical structure for the parameters of a single pulse segment model.*

Parameters	Prior Distributions
k_i	Truncated Poisson Distribution
ω_{k_i}	Uniform Distribution
a_{k_i}	Multivariate Normal Distribution
$\sigma_{k_i}^2$	Jeffrey's uninformative prior
δ_i^2	Inverse Gamma Distribution
Λ_i	Gamma Distribution

Table 6.1: *Prior Distributions for Independent Parameters*

hyperparameter of the amplitudes a_{k_i}.

Most of the priors are based on uninformative Jeffrey's priors and conjugate priors. They have been demonstrated to have a good performance in [33]. The selection of uninformative Jeffrey's priors reduces the dependence of the results on the exact choice of prior. The selection of conjugate priors allows some unknown parameters to be integrated out analytically and thus reduce the computational burden. Table 6.1 shows the parameters of a single pulse segment and their corresponding prior distributions.

In this table, k_i is assigned as a truncated Poisson distribution, conditioned on Λ_i,

which is regarded as the expected number of sinusoids in the ith pulse segment:

$$p(k_i|\Lambda_i) \propto \frac{\Lambda_i^{k_i}}{k_i!}\exp(-\Lambda_i). \tag{6.1}$$

Each frequency value in the vector $\boldsymbol{\omega}_{k_i}$ is assigned as a uniform distribution between 0 and π and all frequencies are assumed i.i.d.. Therefore, the joint prior of frequencies is the product of the prior for each frequency, which is in the form of:

$$p(\boldsymbol{\omega}_{k_i}|k_i) = \left(\frac{1}{\pi}\right)^{k_i}. \tag{6.2}$$

Moreover, the vector of amplitudes $\boldsymbol{a}_{k_i}$ for frequency components is assigned as a multivariate Gaussian distribution with zero mean and covariance matrix, $\sigma_{k_i}^2\boldsymbol{\Sigma}_{k_i}$, as below:

$$p(\boldsymbol{a}_{k_i}|\sigma_{k_i}^2,\delta_i^2) = \frac{1}{|2\pi\sigma_{k_i}^2\boldsymbol{\Sigma}_{k_i}|^{1/2}}\cdot\exp\left[-\frac{\boldsymbol{a}_{k_i}^T\boldsymbol{\Sigma}_{k_i}^{-1}\boldsymbol{a}_{k_i}}{2\sigma_{k_i}^2}\right], \tag{6.3}$$

where $\boldsymbol{\Sigma}_{k_i}^{-1} = \delta_i^{-2}\boldsymbol{D}_i^T\boldsymbol{D}_i$, and δ_i^2 is treated as a random variable. Also, the noise variance in the ith pulse segment, $\sigma_{k_i}^2$, is assigned as a non-informative Jeffrey's prior: $p(\sigma_{k_i}^2|v_i,\gamma_i) \propto 1/\sigma_{k_i}^2$, in which v_i and γ_i are parameters of $p(\sigma_{k_i}^2)$. The last two parameters, Λ_i and δ_i^2, are the hyperparameters of the probability of the frequency numbers $p(k_i|\Lambda_i)$, and the probability of the amplitudes $p(\boldsymbol{a}_{k_i}|\sigma_{k_i}^2,\delta_i^2)$ respectively. Aside from the mathematical convenience, some justifications of these choices of priors have been demonstrated and reported in [162] and the results indicate these non-informative priors perform well for a range of SNRs.

Furthermore, the unknown parameters listed in Table 6.1 are assumed to be independent with each other. Therefore, the joint prior distribution for all parameters in one pulse segment can be considered as a product of each individual prior. If $0 < k_i \leqslant k_{max}$, it can be expressed as:

$$\begin{aligned} p(k_i,\boldsymbol{\theta}_{k_i}) &= p(k_i|\Lambda_i)p(\Lambda_i)p(\boldsymbol{\omega}_{k_i}|k_i)p(\boldsymbol{a}_{k_i}|\sigma_{k_i}^2,\delta_i^2)p(\sigma_{k_i}^2|v_i,\gamma_i)p(\delta_i^2) \\ &\propto \frac{\Lambda_i^{k_i+\epsilon_1-1/2}}{k_i!}\exp[-\Lambda_i(1+\epsilon_2)]\cdot\frac{1}{|2\pi\sigma_{k_i}\boldsymbol{\Sigma}_{k_i}|^{1/2}}\exp\left[-\frac{\boldsymbol{a}_{k_i}^T\boldsymbol{\Sigma}_{k_i}^{-1}\boldsymbol{a}_{k_i}}{2\sigma_{k_i}^2}\right] \\ &\times \left(\frac{1}{\pi}\right)^{k_i}\cdot\frac{\exp[-\gamma_i/(2\sigma_{k_i}^2)]}{\sigma_{k_i}^{2(v_i/2+1)}}\cdot(\delta_i^2)^{-\alpha_{\delta^2}-1}\exp\left[-\frac{\beta_{\delta^2}}{\delta^2}\right]. \end{aligned} \tag{6.4a}$$

where ϵ_1, ϵ_2 are the hyperparameters of Λ_i, and α_{δ^2}, β_{δ^2} are the hyperparameters of δ_i^2.

If $k_i = 0$, it is simplified as:

$$p(\boldsymbol{\theta}_{k_i}|k_i = 0) \propto \exp\left[-\frac{\gamma_i}{2\sigma_{k_i}^2}\right] \cdot \frac{1}{\sigma_{k_i}^{2(v_i/2+1)}} \cdot \exp[-\Lambda_i(1+\epsilon_2)] \cdot \Lambda_i^{\epsilon_1 - 1/2}. \tag{6.4b}$$

6.2.2 Joint posterior distribution for parameters in a single pulse

Based on Bayes's rule, if $0 < k_i \leqslant k_{max}$, by multiplying (4.20) on page 93 and (6.4a), the joint posterior distribution for the unknown parameters can be obtained as:

$$\begin{aligned} p(k_i, \boldsymbol{\theta}_{k_i}|\mathbf{y}_i) &\propto p(k_i, \boldsymbol{\theta}_{k_i}) \cdot p(\mathbf{y}_i|k_i, \boldsymbol{\theta}_{k_i}, T_{2i}, T_{2i-1}) \\ &\propto (2\pi\sigma_{k_i}^2)^{-N_i/2} \cdot \exp\left\{-\frac{\| \mathbf{y}_i - \boldsymbol{D}_i(\boldsymbol{\omega}_{k_i}, T_{2i-1}, T_{2i})\boldsymbol{a}_{k_i} \|^2}{2\sigma_{k_i}^2}\right\} \\ &\times \frac{\Lambda_i^{k_i+\epsilon_1-1/2}}{k_i!} \exp[-\Lambda_i(1+\epsilon_2)] \cdot \frac{1}{|2\pi\sigma_{k_i}\boldsymbol{\Sigma}_{k_i}|^{1/2}} \exp\left[-\frac{\boldsymbol{a}_{k_i}^T\boldsymbol{\Sigma}_{k_i}^{-1}\boldsymbol{a}_{k_i}}{2\sigma_{k_i}^2}\right] \\ &\times \frac{\mathbb{I}_\Omega(k_i, \boldsymbol{\omega}_{k_i})}{\pi^{k_i}} \cdot \frac{\exp[-\gamma_i/(2\sigma_{k_i}^2)]}{\sigma_{k_i}^{2(v_i/2+1)}} \cdot (\delta_i^2)^{-\alpha_{\delta^2}-1} \exp\left[-\frac{\beta_{\delta^2}}{\delta_i^2}\right]. \end{aligned} \tag{6.5}$$

As the amplitudes $\boldsymbol{a}_{k_i}$ and the noise variance $\sigma_{k_i}^2$ are not of most interest, they can be integrated out as nuisance parameters. The detailed derivation of this posterior distribution, including integration of the nuisance parameters, can be found in Appendix B. After the simplification, the joint posterior distribution is given by:

$$\begin{aligned} p(k_i, \boldsymbol{\omega}_{k_i}, \Lambda_i, \delta_i^2|\mathbf{y}_i) &\propto (\gamma_i + \mathbf{y}_i^T\boldsymbol{P}_{k_i}\mathbf{y}_i)^{-(N_i+v_i)/2} \cdot \frac{\Lambda_i^{k_i}}{k_i![(\delta_i^2+1)\pi]^{k_i}} \\ &\times \frac{\Lambda_i^{\epsilon_1-1/2} \cdot e^{-\epsilon_2\Lambda_i}}{e^{\Lambda_i}} \cdot \frac{1}{\delta_i^{2(\alpha_{\delta^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right], \end{aligned} \tag{6.6}$$

where N_i is the length of the ith block of data, $\mathbf{y}_i$, and

$$\boldsymbol{P}_{k_i} = \boldsymbol{I}_{N_i} - \boldsymbol{D}_i\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T \tag{6.7a}$$

$$\boldsymbol{M}_{k_i}^{-1} = \boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1} \tag{6.7b}$$

$$\boldsymbol{m}_{k_i} = \boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i. \tag{6.7c}$$

In the case of $k_i = 0$, $\boldsymbol{a}_0^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{a}_0 \triangleq 0$ and $|2\pi\sigma_{k_i}^2\boldsymbol{\Sigma}_0| \triangleq 1$ are adopted in calculation of the posterior distribution conventionally, as used in [33,34]. The corresponding posterior

distribution is presented as:

$$p(k_i, \boldsymbol{\omega}_{k_i}, \Lambda_i, \delta_i^2 | \mathbf{y}_i, k_i = 0) \propto \frac{1}{\sigma_{k_i}^{N_i+2(v_i/2+1)}} \cdot \exp\left(-\frac{\mathbf{y}_i^T\mathbf{y} + \gamma_i}{2\sigma_{k_i}^2}\right) \times \frac{\Lambda_i^{k_i}/k_i!}{\pi^{k_i}} \cdot \Lambda_i^{\epsilon_1 - 1/2} \cdot \exp[-\Lambda_i(1+\epsilon_2)]. \quad (6.8)$$

6.2.3 Reversible jump MCMC diversity steps

Although the joint posterior distribution in (6.6) has been simplified, it is still highly non-linear in terms of frequencies $\boldsymbol{\omega}_{k_i}$, which means the closed forms of $p(k_i|\mathbf{y}_i)$ and $p(\boldsymbol{\omega}_{k_i}|k_i, \mathbf{y}_i)$ cannot be obtained. Therefore, a sampling technique is required to sample from the complicated joint posterior distribution $p(k_i, \boldsymbol{\omega}_{k_i}|\mathbf{y}_i)$, or from the marginal distributions $p(k_i|\mathbf{y}_i)$ and $p(\boldsymbol{\omega}_{k_i}|k_i, \mathbf{y}_i)$, in order to estimate the parameters of interests.

MCMC techniques are a set of algorithms that are able to sample from probability distributions [100]. A review of the state of the art is presented in Section 5.3. Using these methods, instead of sampling from the probability distribution directly, an ergodic Markov chain is constructed. After a number of iteration steps, the chain will become stable and its equilibrium distribution can be regarded as the desired probability distribution. Also, initial selected values for parameters of interest will converge to their corresponding definite values, which can be regarded as samples from the desired distributions. Having obtained these samples, histograms of these samples can be plotted. Based on the achieved histograms of the samples, a MAP criterion, which has been briefly introduced in Section 4.7.1, is adopted to obtain a mode of the estimated posterior distribution $\hat{p}(k_i|\mathbf{y}_i)$ and $\hat{p}(\boldsymbol{\omega}_{k_i}|k_i, \mathbf{y}_i)$. Then the desired parameters, $(k_i, \boldsymbol{\omega}_{k_i})$, of the ith pulse segment in the model can be easily estimated as:

$$\hat{k}_i|\mathbf{y}_i = \arg\max_{k_i} \hat{p}(k_i|\mathbf{y}_i) \quad (6.9a)$$

$$\hat{\boldsymbol{\omega}}_{k_i}|\hat{k}_i, \mathbf{y}_i = \arg\max_{\boldsymbol{\omega}_{k_i}} \hat{p}(\boldsymbol{\omega}_{k_i}|k_i, \mathbf{y}_i). \quad (6.9b)$$

Moreover, in the ith pulse segment of the signal model, since the number of frequencies k_i is also unknown, a reversible jump technique proposed in [122] is able to estimate the frequency model order k_i. It is able to combine the model order selection and the parameter estimation. This technique allows the proposals to jump between subspaces of different dimensions and visit all relevant model orders. With a certain

ratio that ensures the reversibility and hence the invariance of a Markov chain, samples from the proposal distribution are accepted. Due to this calculated acceptance ratio, the most likely model orders are visited most often and the least likely ones are visited with lower probability. In this way, the computational complexity can be reduced.

In general, there are three simple and commonly used candidate moves for frequency estimation: *birth*, *death* and *update* moves. *Birth* and *death* moves are widely used complementary moves. In a *birth* move, the algorithm proposes a candidate of higher dimension whereas in a *death* move, the algorithm proposes a candidate in the model of lower dimension. If neither move is chosen, an *update* move will be carried out. b_k, d_k and u_k are denoted as the probabilities of each move for frequency estimation. They satisfy the relationship of $b_k + d_k + u_k = 1$ for all k. b_k and d_k use the same definition in (5.38) and (5.39) on page 116 in Section 5.3.5. Detailed descriptions for different rjMCMC diversity steps within one single pulse segment $\mathbf{y}_i$ will be provided in the following sections. In order to simplify the notations, the subscript i in parameters will be omitted in the subsequent descriptions of rjMCMC diversity steps as these steps will perform exact the same for all m pulse segments. [1]

6.2.3.1 *Update* move for a frequency at a constant model order

In this move, the model order of frequencies is fixed and the dimension of frequencies does not change. A hybrid MCMC sampler using MH steps is adopted in this *update* move of frequencies. Within the same dimension, a hybrid mixture kernel can be used as a transition kernel in the MH algorithm, which has been presented in Section 5.3.2.2. One of the advantages of a mixture of kernels lies in the fact that it can incorporate the global proposal to explore vast regions of the state space and the local proposal to discover finer details of the desired target distribution [126, 163]. For example, there are two transition kernels of the Markov chain $\mathcal{K}_1$ and $\mathcal{K}_2$. Both of them have the same invariant distribution $\pi(\cdot)$. In this situation, (5.29) on page 109 can be simplified as a mixture of two independent transition kernels, e.g. $\lambda K_1 + (1-\lambda)K_2$ $(0 \leq \lambda \leq 1)$, which also has the same invariant distribution $\pi(\cdot)$. λ is a factor that tunes the ratio of two transition kernels according to different applications.

In the case of a high-precision frequency estimation, a mixture of kernels can be employed, as suggested in [33]. In the process of frequency estimation, the frequency

[1] N_i, $\mathbf{y}_i$ and $\boldsymbol{D}_i$ will keep the subscript i in order not to make confusion with N, $\mathbf{y}$ and $\mathbf{D}$ in the full length signal model in Section 6.3.

spectrum of a signal is taken as the global proposal distribution, and a random walk perturbation is taken as the local proposal distribution. The invariant distribution is given by $p(\omega_{j,k}|\boldsymbol{\omega}_{-j,k}, \mathbf{y}_i)$, in which $\boldsymbol{\omega}_{-j,k} \triangleq [\omega_{1,k}, \ldots, \omega_{j-1,k}, \omega_{j+1,k}, \ldots]$. The global and the local *update* moves can be described in detail as follows:

1. *Global Update* based on the observations of a signal with a probability of λ. The basic idea is to propose a new frequency, $\omega^*_{j,k}$, independent of the previous one, $\omega_{j,k}$, and to explore regions around obvious peaks in the frequency spectrum of the signal:

$$q_1(\omega^*_{j,k}|\omega_{j,k}) = q_1(\omega^*_{j,k}) \propto \sum_{l=0}^{N_f-1} p_l(\omega^*_{j,k}) \mathbb{I}_{\left[\frac{l\pi}{N_f}, \frac{(l+1)\pi}{N_f}\right]}(\omega^*_{j,k}), \text{ [2]} \tag{6.10a}$$

   where $p_l(\omega^*_{j,k})$ is calculated by the squared modulus of the frequency spectrum of observations $\mathbf{y}_i$ at each individual frequency, $l\pi/N_f$. $N_f \geqslant N_i$ is the number of frequency bins. In [33], the frequency spectrum computed using the DFT is selected as the proposal distribution. However, it sometimes tends to overestimate the number of frequencies if there are a large number of frequency components in a small duration of a signal. Therefore, the multitaper spectrum is chosen in this thesis in estimating the frequency components of signals for its better trade-off between the estimation variance and the frequency resolution. Moreover, the factor $\lambda = 0.2$, tuning the ratio of global distribution and the local distribution, is determined in a heuristic way. It ensures that the corresponding Markov chain has satisfactory properties [33].

2. *Local Update* by random perturbation with a probability of $1 - \lambda$. The proposal distribution is based on a random walk algorithm, which is given by:

$$q_2(\omega^*_{j,k}|\omega_{j,k}) = \mathcal{N}(\omega_{j,k}, \sigma^2_{RW_f}). \tag{6.10b}$$

   σ_{RW_f} is set to $1/(5N_i)$, in which N_i is the length of the ith pulse segment. This selected σ_{RW_f} can ensure a good acceptance probability for a random walk with few parameters, as suggested in [164]. Also, there is a demonstration of good estimation results shown in [33].

Following the MH sampling procedure described in Section 5.3.2.1, according to

[2] $\mathbb{I}_\Omega[\omega]$ equals one if $\omega \in \Omega$ and equals zero if $\omega \notin \Omega$.

Algorithm 6.1: Update move for a frequency

Sample $u_\lambda \sim U_{[0,1]}$;
if $u_\lambda < \lambda$ **then**
  Propose a candidate frequency component according to (6.10a) ;
else
  Propose a candidate frequency component according to (6.10b) ;
end
Evaluate α_{update} based on (6.13) and (6.14) ;
Sample $u \sim U_{[0,1]}$;
if $u \leq \alpha_{update}$ **then**
  The state of the Markov chain becomes $(k, \boldsymbol{\omega}^*_{j,k})$;
else
  The state of the Markov chain remains at $(k, \boldsymbol{\omega}^{(i)}_{j,k})$;
end

(5.23) on page 107, the acceptance ratio of a candidate frequency component can be obtained as:

$$\begin{aligned} r_{update} &\triangleq \text{(posterior distribution ratio)} \times \text{(proposal ratio)} \\ &= \frac{p(k, \boldsymbol{\omega}^*_{j,k}|\mathbf{y}_i)}{p(k, \boldsymbol{\omega}_{j,k}|\mathbf{y}_i)} \times \frac{q(\omega_{j,k}|\omega^*_{j,k})}{q(\omega^*_{j,k}|\omega_{j,k})}. \end{aligned} \tag{6.11}$$

where the vector $\boldsymbol{\omega}^*_{j,k}$ is defined as: $\boldsymbol{\omega}^*_{j,k} \triangleq [\omega_{1,k}, \ldots, \omega_{j-1,k}, \omega^*_{j,k}, \omega_{j+1,k}, \ldots]$. Since

$$p(k, \boldsymbol{\omega}_k|\mathbf{y}_i) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2|\mathbf{y}_i), \tag{6.12}$$

and substituting (6.6) into (6.11), the acceptance ratio r_{update} can be simplified as follows:

$$r_{update} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \boldsymbol{P}^*_k \mathbf{y}_i}\right)^{(N_i+v)/2} \times \frac{q(\omega_{j,k}|\omega^*_{j,k})}{q(\omega^*_{j,k}|\omega_{j,k})}. \tag{6.13}$$

Then the candidate frequency $\omega^*_{j,k}$ is accepted with the following probability according to (5.22) on page 107:

$$\alpha_{update} = \min\{1, r_{update}\}. \tag{6.14}$$

The schema for an *update* move is summarized in Algorithm 6.1.

6.2.3.2 Dimension-changing moves for a frequency

The procedure for a *birth/death* move is different from that for an *update* move because the *birth/death* move contains a change in dimension. In a *birth* move, a candidate in a higher dimension model is proposed, which means, in spectral estimation, a candidate frequency is selected between 0 and π at random following a uniform distribution for simplicity. With a certain acceptance probability, the current state will change from $(k, \boldsymbol{\omega}_k)$ to $(k+1, \boldsymbol{\omega}_{k+1})$. In a *death* move, a candidate is proposed in a lower dimension model and thus an existing frequency component is removed randomly. In order to make the underlying Markov chain reversible with respect to moves across subspaces of different model order, the desired invariant distribution of the chain should be preserved [122]. That is, the probability of moving from model order k to $k+1$ must be equivalent to that of moving from $k+1$ to k. Therefore, when a *death* move is performed, an existing frequency component is randomly selected to be removed such that the state space will change from $(k+1, \boldsymbol{\omega}_{k+1})$ to $(k, \boldsymbol{\omega}_k)$.

Furthermore, both a *birth* move and a *death* move can be generalized to a dimension-change move. The acceptance ratio of the dimension change move can be determined as:

$$r_{change} \triangleq \text{(posterior distribution ratio)} \times \text{(proposal ratio)} \times \text{(Jacobian)}. \tag{6.15}$$

Compared to (5.34) on page 115, the Jacobian term equals to one because the dimension only increases or decreases by one in each iteration.

The first term in (6.15), which is an expression for the posterior ratio within a specific pulse segment, $\mathbf{y}_i$, is presented:

$$\text{posterior distribution ratio} = \frac{p(k+1, \boldsymbol{\omega}_{k+1} | \Lambda, \delta^2, \mathbf{y}_i)}{p(k, \boldsymbol{\omega}_k | \Lambda, \delta^2, \mathbf{y}_i)}. \tag{6.16}$$

Since

$$p(k, \boldsymbol{\omega}_k | \Lambda, \delta^2, \mathbf{y}_i) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2 | \mathbf{y}_i), \tag{6.17}$$

and considering (6.6), the posterior ratio can be obtained:

$$\text{posterior ratio} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \boldsymbol{P}_{k+1} \mathbf{y}_i} \right)^{(N_i+v)/2} \times \frac{1}{(k+1)(\delta^2+1)} \times \frac{\Lambda}{\pi}. \tag{6.18}$$

The second term in (6.15) is the proposal ratio. In a *birth* move, the probability of

choosing the move is b_k, which was defined in (5.38) on page 116. Combined with the probability of proposing a new frequency based on a uniform distribution, the overall proposal distribution for a *birth* move is b_k/π. For the corresponding *death* move, a randomly chosen frequency will be removed from $k+1$ existing frequencies. Since the probability of choosing a *death* move is d_k, which was defined in (5.39) on page 116, and the probability of choosing a specific sinusoid to be removed from $k+1$ frequencies is $1/(k+1)$, the overall probability of a *death* move is $d_{k+1}/(k+1)$. In both (5.38) and (5.39), there is a prior distribution for the number of frequencies, $p(k|\Lambda)$, which was assigned as a Poisson distribution in (6.1). Therefore, the proposal distributions for a *birth* move and a *death* move can be expressed as:

$$q(k+1, \boldsymbol{\omega}_{k+1}|\Lambda, \delta^2, \mathbf{y}_i) = b_k \times \frac{1}{\pi} \tag{6.19}$$

$$q(k, \boldsymbol{\omega}_k|\Lambda, \delta^2, \mathbf{y}_i) = d_{k+1} \div \begin{pmatrix} k+1 \\ 1 \end{pmatrix}. \tag{6.20}$$

As a result, for $k = 0, \ldots, k_{max} - 1$, the proposal ratio can be determined using (6.19) and (6.20):

$$\text{proposal ratio} = \frac{d_{k+1}/(k+1)}{b_k/\pi} = \frac{p(k|\Lambda)}{p(k+1|\Lambda)} \times \frac{\pi}{k+1} = \frac{\pi}{\Lambda}, \tag{6.21}$$

if $k = k_{max}$ (k_{max} is the predefined maximum number of frequency components), only the *death* move is performed; if $k = 0$, only the *birth* move is performed.

Finally, by substituting (6.18) and (6.21) into (6.15), the acceptance ratio of the dimension-change move becomes:

$$r_{change} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \boldsymbol{P}_{k+1} \mathbf{y}_i} \right)^{(N_i+v)/2} \times \frac{1}{(k+1)(\delta^2+1)}. \tag{6.22}$$

Moreover, according to (5.35) on page 115, the acceptance probabilities of the *birth* and the *death* moves will have the following expressions since the *death* move is an inverse process of the *birth* move:

$$\alpha_{birth} = \min\{1, r_{change}\} \tag{6.23a}$$

$$\alpha_{death} = \min\{1, r_{change}^{-1}\}. \tag{6.23b}$$

Algorithm 6.2: Birth move for a frequency

Propose a new frequency component, ω_{k+1}, on $(0, \pi)$, then the new vector of frequencies becomes $\boldsymbol{\omega}_{k+1}^{(i+1)} = [\boldsymbol{\omega}_k^{(i)}, \omega_{k+1}]$;
Evaluate α_{birth} according to (6.22) and (6.23a) ;
Sample $u \sim U_{[0,1]}$;
if $u \leqslant \alpha_{birth}$ **then**
    The state of the Markov chain becomes $(k+1, \boldsymbol{\omega}_{k+1}^{(i+1)})$;
else
    The state of the Markov chain remains at $(k, \boldsymbol{\omega}_k^{(i)})$;
end

Algorithm 6.3: Death move for a frequency

Choose the jth frequency component randomly from the existing frequency vector, then the new vector of frequencies becomes: $\boldsymbol{\omega}_k^{(i+1)} = [\boldsymbol{\omega}_{1:(j-1)}^{(i)}, \boldsymbol{\omega}_{(j+1):(k+1)}^{(i)}]$;
Evaluate α_{death} according to (6.22) and (6.23b) ;
Sample $u \sim U_{[0,1]}$;
if $u \leqslant \alpha_{death}$ **then**
    The state of the Markov chain becomes $(k, \boldsymbol{\omega}_k^{(i+1)})$;
else
    The state of the Markov chain remains at $(k+1, \boldsymbol{\omega}_{k+1}^{(i)})$;
end

The schemas for the *birth* and *death* moves are summarized in Algorithm 6.2 and Algorithm 6.3 respectively.

6.2.3.3 Update of other parameters in a single pulse segment

$\boldsymbol{a}_k$ and σ_k^2 are nuisance parameters and do not need to be updated formally in each iteration of the algorithm. However, they are required in updating one of the hyper-parameters δ^2. In Appendix B, the priors selected for $\boldsymbol{a}_k$ and σ_k^2 are both conjugate priors in order that they can be integrated out from the complicated joint posterior distribution. Using the conjugate priors, the posterior distributions can be put in a much simpler form and the computations can be highly reduced. During the process of integration, the prior and posterior for $\boldsymbol{a}_k$ are both multivariate Gaussian distributions, which are displayed in (6.24a) and (6.24b) respectively. Moreover, the prior and posterior for σ_k^2 are both inverse Gamma distributions, which are displayed in (6.25a)

and (6.25b) respectively.

$$\boldsymbol{a}_k \sim \mathcal{N}(0, \sigma_k^2 \boldsymbol{\Sigma}_k^{-1}) \tag{6.24a}$$

$$\boldsymbol{a}_k|(\mathbf{y}_i, k, \boldsymbol{\omega}_k, \sigma_k^2, \Lambda, \delta^2) \sim \mathcal{N}(\boldsymbol{m}_k, \sigma_k^2 \boldsymbol{M}_k). \tag{6.24b}$$

$$\sigma_k^2 \sim \mathcal{IG}(v, \gamma) \tag{6.25a}$$

$$\sigma_k^2|(\mathbf{y}_i, k, \boldsymbol{\omega}_k, \Lambda, \delta^2) \sim \mathcal{IG}\left(\frac{N_i + v}{2}, \frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k \mathbf{y}_i}{2}\right). \tag{6.25b}$$

In addition to the nuisance parameters, it is necessary to simulate the hyperparameters Λ and δ^2 in each iteration in implementing the estimation algorithm. As suggested in [33, 161], they can be numerically integrated out. In practice, the estimates from $p(\Lambda|\mathbf{y}_i)$ and $p(\delta^2|\mathbf{y}_i)$ are required in implementing the estimation algorithm. In order to simplify the implementation of the algorithm, vague conjugate priors are assigned to Λ and δ^2 as:

$$\Lambda \sim \mathcal{G}a(1/2 + \epsilon_1, \epsilon_2) \tag{6.26a}$$

$$\delta^2 \sim \mathcal{IG}(\alpha_{\delta^2}, \beta_{\delta^2}), \tag{6.26b}$$

where $\epsilon_1 \ll 1$ and $\epsilon_2 \ll 1$ are hyperparameters of Λ. The prior set to Λ is a conjugate prior and the values of ϵ_1 and ϵ_2 are set to close to zero. Moreover, α_{δ^2} and β_{δ^2} are hyperparameters of δ^2. The specification of $\alpha_{\delta^2} = 2$ ensures the mean of the distribution, and $p(\delta^2) > 0$ is only determined by β_{δ^2}. In [33], there is an experimental demonstration that β_{δ^2} has weak influence on the posterior distributions $p(\delta^2|\mathbf{y}_i)$ and $p(k|\mathbf{y}_i)$. For various SNRs and a large range of values of $\beta_{\delta^2} = 1, 10, 100$, the estimated results are stable by experimental observations. Therefore, the procedure of estimation algorithm seems quite insensitive to the hyper-hyperparameters, i.e. ϵ_1, ϵ_2, α_{δ^2} and β_{δ^2}.

When implementing the algorithm, the hyperparameter, Λ, is updated using a MH step on the distribution, $p(\Lambda|k, \boldsymbol{\omega}_k, \delta^2, \mathbf{y}_i)$. It has been derived in [34] that the posterior distribution is in the form of:

$$\begin{aligned} p(\Lambda|k, \mathbf{y}_i) &\propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, |\mathbf{y}_i) \\ &\propto \frac{\Lambda^{\epsilon_1 + k - 1/2} \exp[-\Lambda(1 + \epsilon_2)]}{\Gamma(k + \epsilon_1 - 1/2)} \times (1 + \epsilon_2)^{k + \epsilon_1 + 1/2}, \end{aligned} \tag{6.27a}$$

which gives the update formula as:

$$\Lambda|(k, \mathbf{y}_i) \sim \mathcal{G}a(1/2 + k + \epsilon_1, 1 + \epsilon_2). \tag{6.27b}$$

With respect to the hyperparameter, δ^2, in order not to make the estimation process too complicated, the full conditional distribution $p(\delta^2|k, \boldsymbol{\omega}_k, \Lambda, \boldsymbol{a}_k, \sigma_k^2, \mathbf{y}_i)$ is considered rather than $p(\delta^2|k, \boldsymbol{\omega}_k, \Lambda, \mathbf{y}_i)$. That explains why the update of $\boldsymbol{a}_k$ and σ_k^2 are required in each iteration in the implementation of the estimation algorithm. It also has been derived in [34] that, if $k > 0$, the posterior distribution of δ^2 is expressed as:

$$p(\delta^2|k, \boldsymbol{\omega}_k, \Lambda, \boldsymbol{a}_k, \sigma_k^2, \mathbf{y}_i) \propto \delta^{-2(k+\alpha_{\delta^2}+1)} \times \exp\left[-\frac{1}{\delta^2}\left(\beta_{\delta^2} + \frac{\boldsymbol{a}_k^T \boldsymbol{D}_i^T \boldsymbol{D}_i \boldsymbol{a}_k}{2\sigma_k^2}\right)\right], \tag{6.28a}$$

which indicates the update formula for δ^2:

$$\delta^2|(k, \sigma_k^2, \boldsymbol{a}_k, \boldsymbol{\omega}_k, \mathbf{y}_i) \sim \mathcal{IG}\left(k + \alpha_{\delta^2}, \frac{(\boldsymbol{D}_i \boldsymbol{a}_k)^T (\boldsymbol{D}_i \boldsymbol{a}_k)}{2\sigma_k^2} + \beta_{\delta^2}\right). \tag{6.28b}$$

If $k = 0$, Λ and δ^2 are updated from their prior distributions.

6.3 Temporal estimation in a full length signal model

In Section 3.2, a new *non-parametric* method, termed HTWD-VAD, has been proposed to estimate the pulse locations in the time domain. In this section, a *parametric* method is used and the temporal information, together with the spectral information, are estimated based on a proposed signal model. The change-points, which are related to the pulse locations of a signal, are denoted in the vector of $\mathbf{T}_{2m}$. Since the results from the *non-parametric* methods provide rough estimates of each change-point, and the detected number of change-points are fixed, only the update of change-points need to be performed. Therefore, it can be included in the rjMCMC diversity steps with a fixed dimension. The update procedure is divided into two parts: update of all change-points and update of related parameters, which should be considered in a full length signal model.

6.3.1 Update move for change-points

During the process of the *update* move for change-points, a random walk algorithm is chosen as the proposal distribution due to its simplicity. It means the newly proposed

change-point positions for all pulses depend on their previous values and a local exploration around the initial guesses is performed. The vector of pulse locations, $\mathbf{T}_{2m}$, is a set of integer values. They are updated using a multivariate normal distribution as follows:

$$q_T(\mathbf{T}^*_{2m}|\mathbf{T}_{2m}) = \mathcal{N}(\mathbf{T}_{2m}, \mathbf{\Sigma}_T). \tag{6.29}$$

New candidates $\mathbf{T}^*_{2m}$ are first proposed and then rounded off to the closest integers. The corresponding acceptance ratio r_{update_T} is then evaluated. Via (5.34) on page 115, the Jacobian term can be omitted as the dimension of change-points does not change. The acceptance ratio is expressed as:

$$r_{updateT} = \text{(posterior distribution ratio)} \times \text{(proposal ratio)}. \tag{6.30}$$

According to (6.30), the first term, the posterior distribution ratio needs to be considered in the full length signal model. The joint prior distribution for all parameters should also be studied in the full length model, $\mathbf{y}$, which can be expressed as:

$$\begin{aligned}
p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \mathbf{\Psi}_m) &= p(\mathbf{T}_{2m})p(\boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \mathbf{a}_{k_m}|\sigma_n^2)p(\sigma_n^2) \\
&\propto \frac{1}{|2\pi\sigma_n^2\mathbf{\Sigma}_{k_m}|^{1/2}} \cdot \exp\left[-\frac{\mathbf{a}_{k_m}\mathbf{\Sigma}_{k_m}^{-1}\mathbf{a}_{k_m}}{2\sigma_n^2}\right] \times \frac{1}{\sigma_n^{2(v/2+1)}} \cdot \exp\left[-\frac{\gamma}{2\sigma_n^2}\right] \\
&\times \prod_{i=1}^{m}\left\{\left(\frac{1}{\pi^{k_i}}\right) \cdot \frac{\Lambda_i^{k_i}}{k_i!} \cdot e^{-\Lambda_i} \times \Lambda_i^{\epsilon_1 - 1/2} \cdot \exp[-\epsilon_2\Lambda_i]\right\} \\
&\times \prod_{i=1}^{m}\left\{\frac{1}{\delta_i^{2(\alpha_{\delta_i}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right]\right\} \\
&\times \left(\frac{1}{N-1} \cdot \frac{1}{N-2} \cdots \frac{1}{N-2m}\right),
\end{aligned} \tag{6.31}$$

where most of the parameters have the same definitions in Section 4.5.1. Moreover, the joint prior has a similar expression to that in (6.4a). The set of parameters in each pulse segment are assumed to be independent with each other. Furthermore, σ_n^2 and $\mathbf{a}_{k_m}$ represent the noise variance and the amplitudes of frequencies in the full length of signal model, $\mathbf{y}$, respectively. According to the selection of parameters in a specific pulse segment, the similar priors for σ_n^2 and $\mathbf{a}_{k_m}$ can be given as:

$$\sigma_n^2 \sim \mathcal{IG}(v/2, \gamma/2) \tag{6.32a}$$

$$\mathbf{a}_{k_m} \sim \mathcal{N}(0, \sigma_n^2\mathbf{\Sigma}_{k_m}^{-1}), \tag{6.32b}$$

where

$$\boldsymbol{\Sigma}_{k_m}^{-1} \triangleq \begin{bmatrix} \boldsymbol{\Sigma}_{k_1}^{-1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_{k_2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\Sigma}_{k_m}^{-1} \end{bmatrix},$$

and $\boldsymbol{\Sigma}_{k_i}^{-1} = \delta_{k_i}^{-2} \cdot \boldsymbol{D}_i^T \boldsymbol{D}_i$, in which $i = 1, \ldots, m$. By multiplying (4.18) on page 92 and (6.31), the posterior distribution for $\mathbf{T}_{2m}$ can be obtained based on Bayes's rule. The derivation is similar to the update of parameters in a single pulse segment. Detailed derivation and the simplification of the joint posterior distribution can be found in Appendix B. Since

$$p(\mathbf{T}_{2m}|\boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2, \mathbf{y}) \propto p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2|\mathbf{y}), \tag{6.33}$$

the posterior distribution for change-points $\mathbf{T}_{2m}$ can be represented as follows:

$$\begin{aligned} & p(\mathbf{T}_{2m}|\boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2, \mathbf{y}) \\ & \qquad \propto (\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m} \mathbf{y})^{-(N+v)/2} \\ & \qquad \times \prod_{i=1}^{m} \left\{ \left(\frac{1}{1+\delta_i^2} \right)^{2k_i} \delta_i^{-2(\alpha_{\delta_i^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2} \right] \right\} \\ & \qquad \times \prod_{i=1}^{m} \left\{ \frac{1}{k_i!} \cdot \left(\frac{\Lambda_i}{\pi} \right)^{k_i} \times \Lambda_i^{\epsilon_1 - 1/2} \cdot e^{-\Lambda_i(1+\epsilon_2)} \right\}, \end{aligned} \tag{6.34}$$

where

$$\boldsymbol{P}_{k_m} = \boldsymbol{I}_N - \mathbf{D}\boldsymbol{M}_{k_m}\mathbf{D}^T \tag{6.35a}$$

$$\boldsymbol{M}_{k_m}^{-1} = \mathbf{D}^T\mathbf{D} + \boldsymbol{\Sigma}_{k_m}^{-1} \tag{6.35b}$$

$$\boldsymbol{m}_{k_m} = \boldsymbol{M}_{k_m}\mathbf{D}^T\mathbf{y}. \tag{6.35c}$$

These definitions of $\boldsymbol{P}_{k_m}$, $\boldsymbol{M}_{k_m}^{-1}$ and $\boldsymbol{m}_{k_m}$ are similar to those defined in (6.7) on page 126, which are considered in a single pulse segment of the signal model.

Moreover, the second term in (6.30), i.e. the proposal ratio can be obtained based on the random walk algorithm shown in (6.29). Therefore, substituting (6.29) and (6.34)

into (6.30), the acceptance ratio of *update* move for change-points, $\boldsymbol{T}_{2m}$, becomes:

$$r_{updateT} = \left(\frac{\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m} \mathbf{y}}{\gamma + \mathbf{y}^T \boldsymbol{P}^*_{k_m} \mathbf{y}}\right)^{(N+v)/2} \times \frac{q_T(\mathbf{T}_{2m}|\mathbf{T}^*_{2m})}{q_T(\mathbf{T}^*_{2m}|\mathbf{T}_{2m})}. \tag{6.36}$$

Also, the corresponding acceptance probability for all change-points can be calculated using (5.35).

$$\alpha_{updateT} = \min\{1, r_{updateT}\}. \tag{6.37}$$

Finally, having obtained the posterior distribution for $\mathbf{T}_{2m}$, an MAP criterion can be used to achieve the desired estimates of change-points.

$$\hat{\mathbf{T}}_{2m}|\mathbf{y} = \underset{\mathbf{T}_{2m}}{\arg\max} \quad \hat{p}(\mathbf{T}_{2m}|\boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}^2_{1:m}, \mathbf{y}). \tag{6.38}$$

6.3.2 Update of related parameters

The update of change-points $\mathbf{T}_{2m}$ in the full length signal model, $\mathbf{y}$, requires updates of $\mathbf{a}_{k_m}$ and σ_n^2. According to the derivation in Appendix B, the posterior distributions of $\mathbf{a}_{k_m}$ and σ_n^2 can be integrated out and put in standard forms:

$$\sigma_n^2|\mathbf{y} \sim \mathcal{IG}\left(\frac{N+v}{2}, \frac{\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m} \mathbf{y}}{2}\right) \tag{6.39}$$

$$\mathbf{a}_{k_m}|\mathbf{y} \sim \mathcal{N}(\boldsymbol{m}_{k_m}, \sigma_n^2 \boldsymbol{M}_{k_m}), \tag{6.40}$$

where $\boldsymbol{P}_{k_m}$, $\boldsymbol{M}_{k_m}$ and $\boldsymbol{m}_{k_m}$ have been already defined in (6.35). Moreover, when updating $\boldsymbol{M}_{k_m}$, each single element, δ_i^2, in $\boldsymbol{\delta}^2_{1:m}$ is updated in the corresponding pulse segment, $\mathbf{y}_i$ using (6.28b). In this way, $\boldsymbol{\delta}^2_{1:m} = (\delta_1^2, \ldots, \delta_m^2)$ can be updated in each iteration in the estimation algorithm.

6.3.3 Procedure of the update move for pulse locations

The schema for the *update* move for change-points, $\mathbf{T}_{2m}$, is described in Algorithm 6.4. Using this *update* move for the vector of change-points in the full length signal model, $\mathbf{y}$, the estimates of pulse locations from the *non-parametric* estimation method can be refined using a *parametric* method.

Algorithm 6.4: Update move for change-points $\mathbf{T}_{2m}$

Propose a new candidate vector for pulse locations $\mathbf{T}_{2m}$ according to (6.29) ;
Evaluate α_{update_T} according to (6.36) and (6.37) ;
Sample $u \sim U_{[0,1]}$;
if $u \leqslant \alpha_{update}$ **then**
    The state of the Markov chain for vector of pulse locations becomes $\mathbf{T}^*_{2m}$;
else
    The state of the Markov chain remains at $\mathbf{T}_{2m}$;
end
Sample σ_n^2 and $\mathbf{a}_{k_m}$ according to (6.39) and (6.40) respectively ;

6.4 Algorithm for the joint estimation system

In all, the procedure of a estimation system for measured ultrasound echo signals can be summarized in Algorithm 6.5. The estimation system combines the *non-parametric* method and the *parametric* method so as to achieve better estimates of parameters in the proposed signal model, particularly its temporal and spectral contents, which are of most interest.

6.5 Evaluation of the parametric estimation method

An extensive Monte Carlo study has been carried out to check the stability and accuracy of the proposed estimation algorithm, since the probabilities under investigation are not available in closed forms. Moreover, the simulated signal with ground truths used in Chapter 3 is used in this section to compare the *non-parametric* method and the *parametric* method.

The construction of the simulated signal and its parameter setup are described in Appendix A. There are two pulse segments in the signal, with locations at (450, 600) and (750, 850) respectively, in the total length of 1500 data points. The synthetic signal is shown in Figure 3.6(a) on page 40. The estimation results obtained from the *non-parametric* method have been discussed in Chapter 3, and will be compared to the results from the *parametric* method in Section 6.5.3. With respect to spectral content of the simulated signal, there are two frequencies in the first pulse. In the second pulse, there are three frequencies, two of which are closely-spaced compared to the width of the pulse. However, when the *non-parametric* methods were used in Section 3.3.5.1, the frequency components 0.3π and 0.32π merged into a single wide peak, and thus neither the DFT spectrum displayed in Figure 3.17 on page 62 nor the

Algorithm 6.5: Algorithm for the estimation system

***Non-parametric* estimation (coarse estimation) ;**
Coarse estimation for pulse locations $\boldsymbol{T}_{init}$;
Coarse estimation for frequency components $\boldsymbol{\omega}_{k_m}$;
***Parametric* estimation using a rjMCMC algorithm (estimation refinement) ;**
Initialization: set $\{k, \boldsymbol{\theta}_k\}_m^{(0)} = \{\boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \mathbf{a}_{k_m}, \boldsymbol{\sigma}_n^2\}$, $\boldsymbol{T}_{2m}^{(0)} = \boldsymbol{T}_{init}$;
Iteration: ;
for $iter = 1$ **to** $numIteration$ **do**
    Update the change-point vector, $\boldsymbol{T}_{2m}^{(iter)}$, using random walk algorithm according to Algorithm 6.4 ;
    for $i = 1$ **to** m **do**
        a. Sample hyperparameters Λ_i and δ_i^2 ;
        b. Sample $u \sim U_{[0,1]}$;
        if $u \leq b_{k_i^{(iter)}}$ **then**
            Perform *birth* move of a new frequency according to Algorithm 6.2 on page 133 ;
        else if $(u \leq b_{k_i^{(iter)}} + d_{k_i^{(iter)}})$ **then**
            Perform *death* move of an existing frequency according to Algorithm 6.3 on page 133 ;
        else
            Perform *update* move of a frequency randomly according to Algorithm 6.1 on page 130 ;
        end
        c. Sample nuisance parameters $\boldsymbol{a}_{k_i}$ and $\sigma_{k_i}^2$;
    end
end

multitaper spectrum displayed in Figure 3.18 on page 63 were able to discriminate them due to their limited frequency resolutions. However, if a *parametric* method is used, the frequency resolution mainly depends on the SNR and will not be only limited by the length of pulses.

6.5.1 Parameter setup for the estimation algorithm

There are some parameter specifications that are required in the developed estimation algorithm. These parameters have no big influence on the estimation results and are listed in Table 6.2. For the ith pulse segment, v_i and γ_i are parameters of the noise variance $\sigma_{k_i}^2$. They are both set to zero in order to ensure that an uninformative Jeffrey's prior is selected, i.e. $p(\sigma_{k_i}^2) \propto 1/\sigma_{k_i}^2$. Moreover, Λ_i and δ_i^2 are hyperparameters of the number of frequency k_i; ϵ_1 and ϵ_2 are hyperparameters of Λ_i; $\alpha_{\delta_i^2}$ and $\beta_{\delta_i^2}$ are hyperpa-

Parameters	v_i	γ_i	ϵ_1	ϵ_2	$\alpha_{\delta_i^2}$	$\beta_{\delta_i^2}$	λ	σ_{RW}	σ_T
Values	0	0	0.01	0.01	2	10	0.2	$1/(5N_i)$	5

Table 6.2: *Parameters of the algorithm*

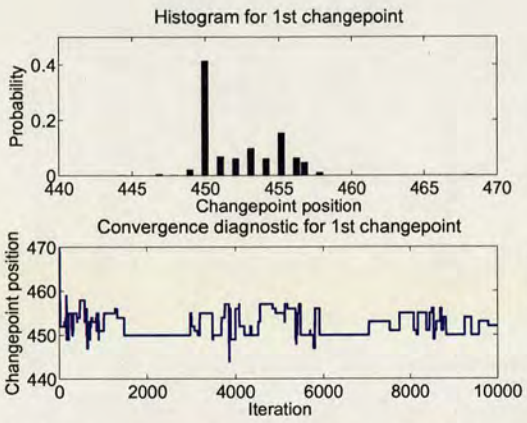
rameters of δ_i^2. Therefore, ϵ_1, ϵ_2, $\alpha_{\delta_i^2}$ and $\beta_{\delta_i^2}$ are hyper-hyperparameters of the signal model. In [33], it is demonstrated that the estimation procedure is very insensitive to the specification of the hyper-hyperparameters when it is applied to frequency estimation. They are assigned the same values that are used in [33]. Other parameters, such as λ, σ_{RW} and σ_T, are chosen in a rather heuristic way. $\lambda = 0.2$ and $\sigma_{RW} = 1/(5N_i)$, where N_i is the length of the ith pulse segment, are used in [33] and achieve good estimated results. Therefore, they are also adopted in this developed estimation algorithm. Furthermore, σ_T is a parameter in the full length signal model, which is used for updating the pulse locations. The value $\sigma_T = 5$ is the first value that is tried, and it provides good results of estimation.

6.5.2 Results for a single realisation of a simulated signal

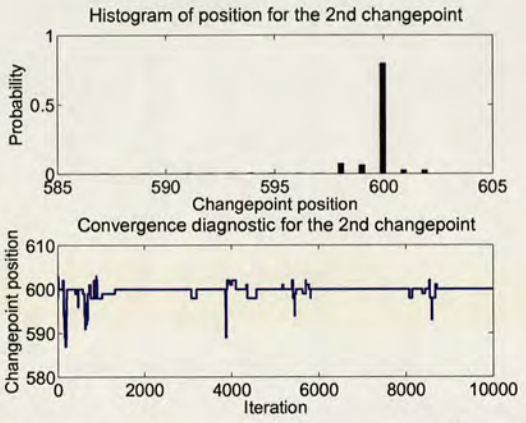
In this section, only one realisation of the simulated signal embedded in Gaussian noise as shown in Figure 3.6(a) on page 40 is taken to explain how the algorithm performs and how the estimates are obtained. The ground truths of model parameters are given in Appendix A. In Section 6.5.3, a Monte Carlo run of one hundred realizations will be performed to evaluate this developed algorithm.

The number of iterations was set to 10000, which was shown to be sufficient because the instantaneously estimated number of frequencies for two pulse segments, $\hat{p}(k_1|\mathbf{y}_1)$ and $\hat{p}(k_2|\mathbf{y}_2)$ shown in Figure 6.4, were stabilized. After a certain number of iterations, i.e. burn-in periods which is defined in Section 5.3.1, all the probabilities of parameters reach their stationary states. Then the MAP criterion is used to produce the estimated model parameters.

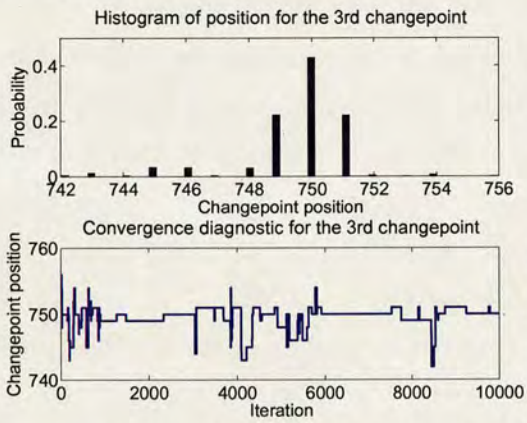
Figure 6.3 shows the estimated pulse locations for each change-point of the two pulse segments respectively. There are four subfigures and each has two panels. The top panel displays a histogram of the posterior probability of each estimated change-point, e.g. $\hat{p}(T_i|\mathbf{y})$. The bottom panel displays the instantaneous estimation of the corresponding $\hat{p}(T_i|\mathbf{y})$. As the coarse estimation obtained from the *non-parametric* method render good initial guesses, the burn-in periods are not long, implying that the Markov



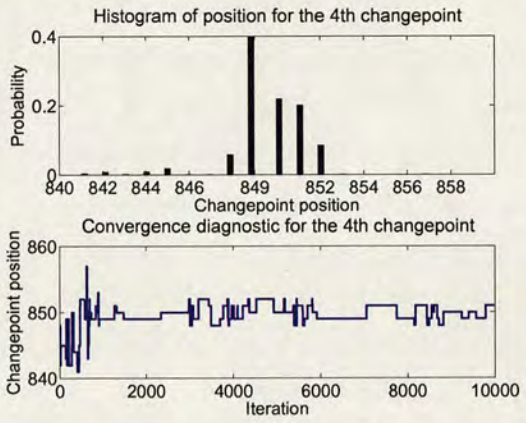
(a) Position estimation refinement for the 1^{st} change-point.



(b) Position estimation refinement for the 2^{nd} change-point.



(c) Position estimation refinement for the 3^{rd} change-point.

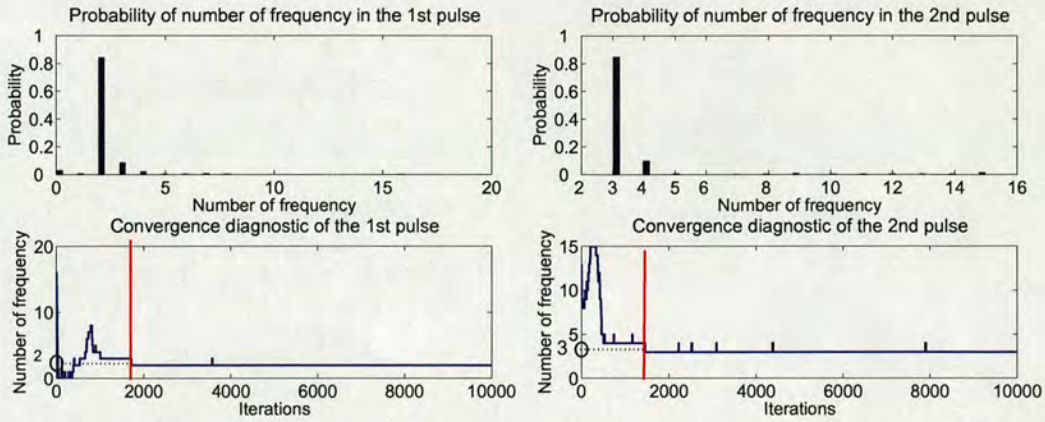


(d) Position estimation refinement for the 4^{th} change-point.

Figure 6.3: *Estimation of positions for all four change-points of two pulses in the simulated signal.*

chain converges quickly. Furthermore, by throwing away the samples in the burn-in periods, the rest of samples are used to plot the histograms of the posterior probabilities.

In terms of the frequency estimation for each pulse segment, the estimates of the number of frequencies are illustrated in Figure 6.4. There are also two panels in each subfigure. Similar to the pulse location estimation, the histograms of frequency numbers are displayed at the top panels and the corresponding instantaneous estimations are displayed at the bottom panels. It can be seen that the burn-in periods are longer than those for pulse location estimation as the initial guesses for the number of frequencies are chosen randomly. Another reason of having longer burn-in periods is



(a) Number of frequency estimation for the 1^{st} pulse.

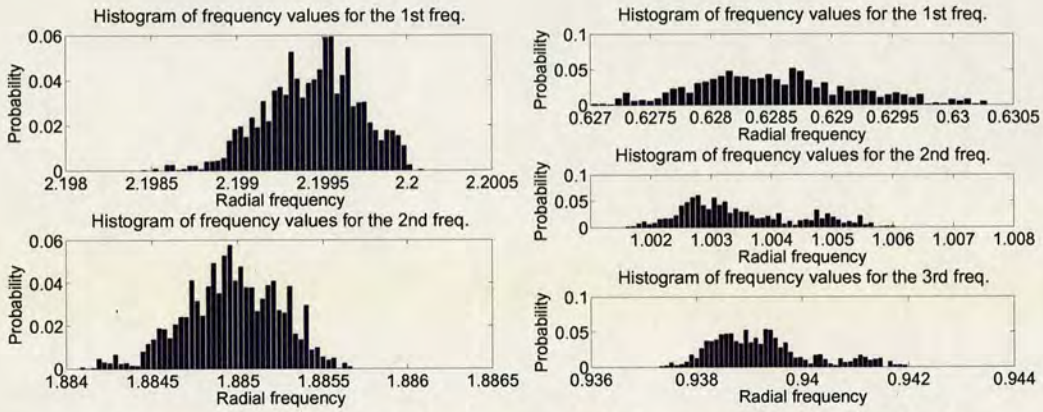
(b) Number of frequency estimation for the 2^{nd} pulse.

Figure 6.4: *Number of frequency estimation and the convergence diagnostics.*

that only the estimates of pulse locations have converged to their true values, the estimates of frequency components in each pulse segment are able to converge. In addition to estimation of the number of frequencies, the histograms of the posterior distributions of frequency values are of equal importance in the estimation. By discarding the values from the first 5000 iterations, the remaining 5000 values are plotted in histograms and are displayed in Figure 6.5. The estimated frequency values can be obtained by using the MAP criterion. It is worth noticing from the estimation results that three frequency components are detected in the second pulse segment. The two closely-spaced frequencies have been discriminated. Therefore, by using the proposed *parametric* method, the third spectral component can be detected and its value can be estimated in the second pulse of the simulated signal, which demonstrates that the frequency resolution is improved.

6.5.3 Evaluation using one hundred realisations

In order to evaluate the performance of the developed estimation algorithm, a Monte Carlo method is employed when a stochastic model is assumed. Unlike the point estimation in a deterministic way, the process of the Monte Carlo method is repeated thousands of times with a set of random variables. Finally, a distribution of outcomes shows not only what the most likely estimate is, but also what ranges are plausible [100]. Accordingly, an extensive Monte Carlo study of the performance of the algorithm is presented in this section. Also, the estimation results are compared to



(a) Frequency values in the 1^{st} pulse.

(b) Frequency values in the 2^{nd} pulse.

Figure 6.5: *Histograms of frequency values estimation for the two pulses.*

those obtained from the *non-parametric* methods.

The algorithm is applied to the simulated signal with 100 different noise realisations, which is termed as 100 Monte Carlo runs. In each Monte Carlo run, 10000 iterations are set to ensure the Markov chain converges. Take a specific parameter of interest as an example, by throwing away values of the first 5000 iterations, the remaining 5000 values are regarded as samples drawn from the posterior distribution of the parameter. Having obtained 100 estimates of a specific parameter from 100 Monte Carlo runs, its mean and SD value can be achieved.

Table 6.3 compares the estimation results of the change-points in the simulated signal using both the *non-parametric* method and the *parametric* method. It demonstrates that the means of the *parametric* method are closer to the ground truth; and the SDs of the *parametric* method are smaller than those of the *non-parametric* method. This indicates the superiority of the proposed estimation algorithm using the *parametric* method with higher accuracy and better stability.

Figure 6.6 illustrates the histograms of the estimated number of frequencies for two pulse segments in the simulated signal with 100 Monte Carlo runs. In each run, estimates for the posterior probabilities of model order $\hat{p}^{(5000)}(K_1 = k_1|\mathbf{y}_1)$ and $\hat{p}^{(5000)}(K_2 = k_2|\mathbf{y}_2)$ can be obtained using the MAP criterion. Then the obtained estimates from 100 monte Carlo runs are used to form the displayed histograms. In Figure 6.6, the means and the SDs of the estimated posterior probability of the number of frequencies are illustrated. The green bars denote the mean values of the estimated number of frequencies; the blue and red bars denote the $mean - SD$ and $mean + SD$ of the estimates

Change-point	Ground Truth	*non-parametric* method (mean $\pm$ SD)	*parametric* method (mean $\pm$ SD)
1^{st} change-point	450	460 ± 1	452 ± 2
2^{nd} change-point	600	602 ± 4	600 ± 0
3^{rd} change-point	750	760 ± 2	750 ± 0
4^{th} change-point	850	850 ± 1	850 ± 1

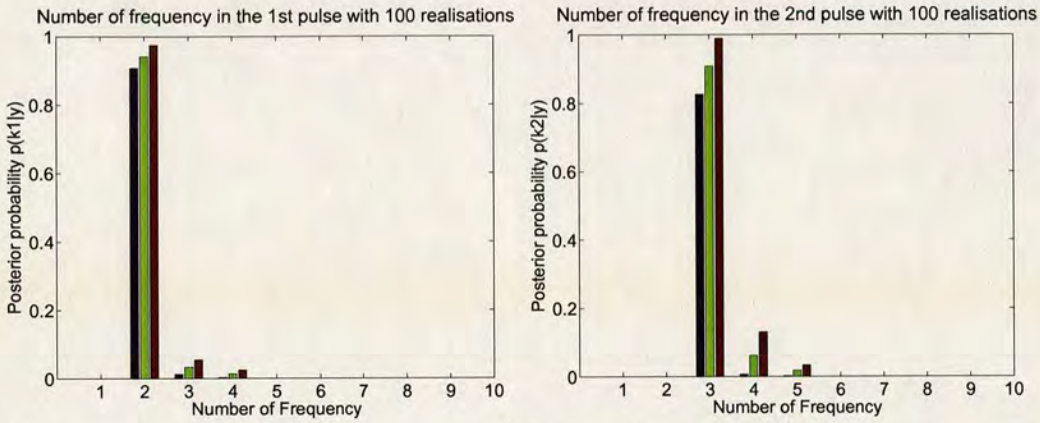
Table 6.3: *Comparison of the estimation of pulse locations using the non-parametric and the parametric method.*

respectively. The peaks in all three statistics at $k_1 = 2$ in Figure 6.6(a) and $k_2 = 3$ in Figure 6.6(b) indicate good estimation performance of the developed algorithm.

Moreover, the estimated frequency values obtained from 100 Monte Carlo runs for two pulse segments of the simulated signal are compared to the ground truths, and are also compared to the results using the *non-parametric* estimation method using multitaper estimates. Table 6.4(a) provides the comparison for the first pulse segment. Compared with the ground truths, both the *non-parametric* method and the *parametric* method provide good mean estimates. However, in terms of the SDs, the *parametric* method outperforms the *non-parametric* method with much smaller variations. Table 6.4(b) gives the comparison for the second pulse segment. The *parametric* method not only provides more accurate mean estimates and smaller SDs, but can distinguish the two closely-spaced frequencies as well. In contrast, the *non-parametric* method regards the two closely-spaced frequencies as one single frequency, which signifies its poor performance in frequency resolution.

6.5.4 Reconstruction of the simulated signal

In Section 4.2, there are four steps addressed in the process of signal model building before a model is applied for real world applications. The first three steps have been discussed in previous sections. The last step is evaluating the performance of a candidate model. It consists of evaluating the performance of the estimation algorithm for model parameters, and validating the proposed model to check how well the signal model fits the observed data. In Section 6.5.2 and Section 6.5.3, the performance of the developed estimation algorithm for model parameters have been investigated in detail. However, model validation has not been discussed yet.



(a) Histogram showing the mean, $mean - SD$, $mean + SD$ of the estimate of $p(k_1|\mathbf{y}_1)$ for the first pulse.

(b) Histogram showing the mean, $mean - SD$, $mean + SD$ of the estimate of $p(k_2|\mathbf{y}_2)$ for the second pulse.

Figure 6.6: *Histograms of number of frequencies using* 100 *noise realizations.*

Freqs.	Ground Truth	*non-parametric* method (mean $\pm$ SD)	*parametric* method (mean $\pm$ SD)
1^{st} freq.	2.1991	2.2000 ± 0.0364	$2.1995 \pm 4.3051 \times 10^{(-4)}$
2^{nd} freq.	1.8849	1.8870 ± 0.0530	$1.8850 \pm 7.2931 \times 10^{(-4)}$

(a) Frequency value comparison for the 1^{st} pulse segment

Freqs.	Ground Truth	*non-parametric* method (mean $\pm$ SD)	*parametric* method (mean $\pm$ SD)
1^{st} freq.	0.6283	0.6200 ± 0.2097	$0.6285 \pm 4.3019 \times 10^{(-4)}$
2^{nd} freq.	0.9425	0.9790 ± 0.2196	$0.9387 \pm 4.8732 \times 10^{(-4)}$
3^{rd} freq.	1.0053	0.0	$1.0027 \pm 5.3047 \times 10^{(-4)}$

(b) Frequency value comparison for the 2^{nd} pulse segment

Table 6.4: *Comparison of frequency values between the non-parametric method and parametric method for two pulse segments.*

Generally, the objective of model validation is to test whether the model satisfies the requirements of the intended application, e.g. some certain criteria that specify the performance of the model. Specifically, the goal of model validation can be described as whether the model sufficiently agrees with the observations, whether the model can characterise the actual signal generation system, or whether the model is able to solve the problem that originates the design process. Most of the investigations into

model validation concentrate on the existence of the mismatch between the model and the observed data, or the residual process. They are calculated in a least-square error statistical sense. One of the key validation techniques focuses on whether the residual process approximate random errors. If the residuals behave randomly, it suggests that the model fits the data well. There are a number of statistical techniques available for checking whether a process is a realisation of white noise [51]. For instance, autocorrelation test, partial correlation test and power spectrum density test. Furthermore, there is another method known as cross-validation to demonstrate whether the model is sufficiently accurate for the purpose. The basic idea of this method is that one set of data is used to fit the model and another statistically independent set of data is used to test the model.

Model validation is a very popular research topic in signal model building process. There are numerous techniques available for validating the proposed signal model and finally choosing the most appropriate model among the candidates. However, in this thesis, a sum of sinusoidal model is proposed because the transmit pulse waveform is composed of a six-period sinusoidal signal. The focus is put on estimating model parameters rather than choosing an appropriate signal model. Therefore, the model validation techniques will not be expanded, and only a simple way of checking whether the estimated parameters are sufficiently accurate is used.

6.5.4.1 Estimation of amplitudes

In the proposed signal model, amplitudes for frequency components $\mathbf{a}_{k_m}$ are not of most interest in the analysis of measured ultrasound echo signals. However, they are required in reconstructing a signal. With the obtained reconstructed signal, the estimation accuracy can be checked by comparing it with the original synthetic signal.

The simulated signal used in evaluating the estimation performance in Section 6.5.3 is used in this section to study the estimation of amplitudes for frequencies. The proposed estimation algorithm is applied to the simulated signal for 100 times with different noise realisations using 100 Monte Carlo runs. In each single realisation, the Markov chain reaches its equilibrium distribution, and the values in the first 5000 iterations are discarded and the rest of them are regarded as samples from the desired posterior distribution of the amplitudes $p(\boldsymbol{a}_{k_i}|\mathbf{y}_i)$. They can be updated according to (6.24b) on page 134.

Using these samples, a histogram is plotted and an MAP estimates for amplitudes

Ground Truth $\{a_k\}_1$	0	4.4721	3.1623	3.1623
Estimated $\{\hat{a}_k\}_1$	0.0124	4.4641	3.0956	3.2150
(mean $\pm$ SD)	$\pm$0.1172	$\pm$0.0491	$\pm$0.0603	$\pm$0.0607

(a) Estimated results for amplitudes for the 1^{st} pulse

Ground Truth $\{a_k\}_2$	3.1623	5.4772	3.7175	5.1167	3.1623	5.4772
Estimated $\{\hat{a}_k\}_2$	3.1482	5.4684	3.7553	5.0952	3.1087	5.5071
(mean $\pm$ SD)	$\pm$0.2041	$\pm$0.1544	$\pm$0.4438	$\pm$0.2080	$\pm$0.3532	$\pm$0.3400

(b) Estimated results for amplitudes for the 2^{nd} pulse

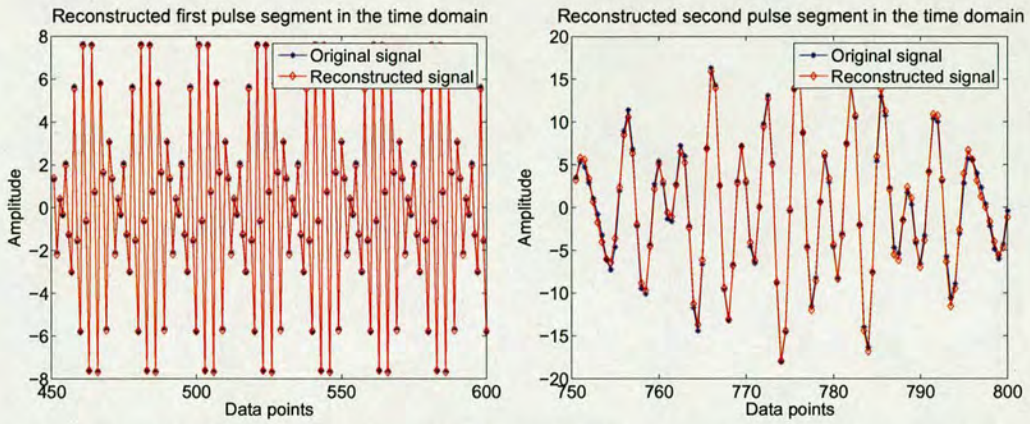
Table 6.5: *Estimated results for amplitudes for two pulses using* 100 *realizations*

$\hat{\boldsymbol{a}}_{k_i}|\mathbf{y}_i = \arg\max_{\boldsymbol{a}_{k_i}} \hat{p}(\boldsymbol{a}_{k_i}|\mathbf{y}_i)$ are presented. Moreover, 100 estimates for each amplitude a_{c,k_i} or a_{s,k_i}, $(i = 1, 2)$ are averaged to provide a mean and a SD value of each amplitude for a single frequency component. Table 6.5(a) and Table 6.5(b) display the resulting estimated amplitudes for the first pulse and the second pulse segments respectively. Compared to the estimated means and SDs of frequency values, the mean errors are larger and the SD values are higher for the estimated amplitudes. This indicates that the developed estimation algorithm provides better estimates for frequency values than the corresponding amplitude values.

6.5.4.2 Signal reconstruction

In the proposed signal model, the start and end points of each pulse segment, the number of frequencies, the frequency values and their corresponding amplitudes of frequencies in each pulse are the model parameters that are necessary to reproduce the original synthetic signal. Apart from the algorithm evaluation addressed in Section 6.5.2 and Section 6.5.3, an alternative way to quantify the performance of the algorithm can be made by looking at the overall reconstruction of the underlying noise-free signal as the noise realisations are different in 100 repeats of the estimation algorithm.

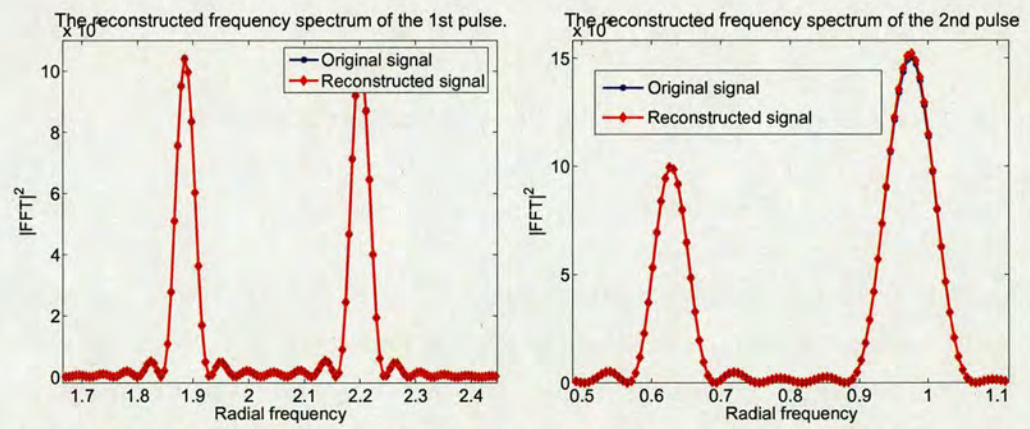
Figure 6.7 illustrate the reconstruction of the 1^{st} pulse and the 2^{nd} pulse of the original signal in the time domain, which is compared to the underlying noise-free signal, s. The blue star-marked line denotes the original signal segment and the red diamond-marked line denotes the reproduced signal segment. They are in good agreement in both Figure 6.7(a) and Figure 6.7(b), implying good estimation of parameters for two pulse segments in the simulated signal. In the frequency domain, the squared mod-



(a) Reconstruction of the 1^{st} pulse segment in the signal.

(b) Reconstruction of the 2^{nd} pulse segment in the signal.

Figure 6.7: *Reconstruction of the pulse segments in the signal without noise.*



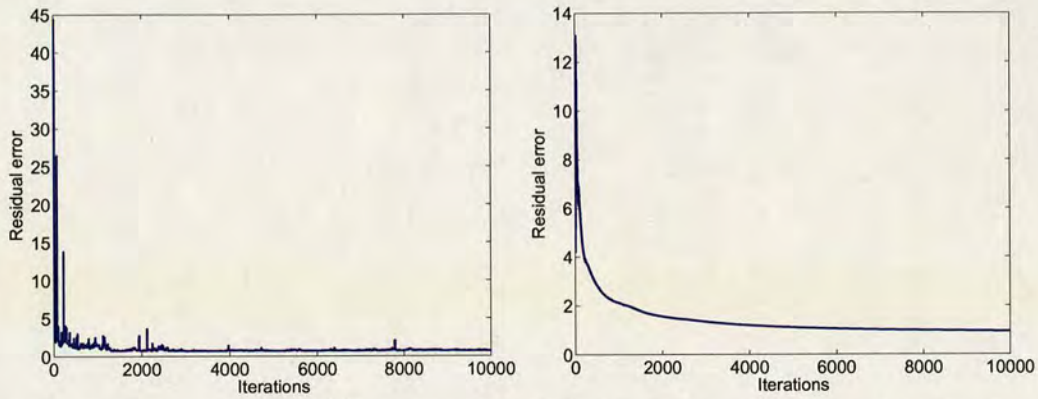
(a) $|FT|^2$ of the 1^{st} pulse segment in the signal.

(b) $|FT|^2$ of the 2^{nd} pulse segment in the signal.

Figure 6.8: *Squared modulus of the FT of the original signal and the reconstructed signal.*

ulus of the FT of the original signal and the reconstructed signal are compared and plotted in Figure 6.8(a) and Figure 6.8(b) for two pulse segments respectively. A good fit in the frequency domain can also be seen for each pulse segment.

Furthermore, two plots of the residual errors of the signal reconstructions, as a function of the number of iterations are shown in Figure 6.9. Since there are 100 runs of the estimation algorithm and each run of the algorithm produces a sequence of residual errors with 10000 iterations, there are 100 sequences of residual errors obtained.



(a) Residual error against the number of iterations based on the current iteration.

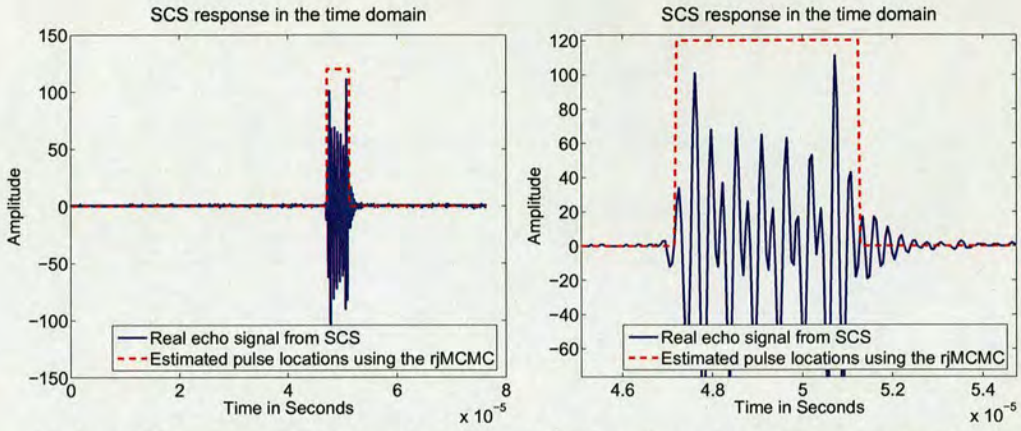
(b) Residual error against the number of iterations based on all preceding iterations.

Figure 6.9: *Plot of the residual errors against the number of iterations.*

For a single run of the algorithm, the residual error can be calculated as:

$$\text{Residual Error}: \quad \mathbf{e} \triangleq \| \mathbf{s}_r^{(iter)} - \mathbf{s} \| /N, \tag{6.41}$$

where N is the length of signal in data points, $\mathbf{s}_r^{(iter)}$ represents the reconstructed signal at the current iteration, $\mathbf{s}$ denotes the original underlying noise-free signal and $\| \cdot \|$ represents Euclidean norm operation. The same procedure is performed on 100 sequences of the residual errors obtained from 100 repeats of the developed estimation algorithm. They are finally averaged to provide an averaged sequence of errors $\mathbf{e}^{(avg)} = (1/100) \sum_{l=1}^{100} \mathbf{e}^{(l)}$. In Figure 6.9(a), the residual errors are computed using the reconstructed signal obtained from the current iteration. However, in Figure 6.9(b), the residual errors are computed using the reconstructed signal based on all previous iterations rather than the current iteration only. In this case, the reconstruction at the iteration, *iter*, becomes: $\mathbf{s}_r^{(iter)} = (1/iter) \sum_{v=1}^{iter} \mathbf{s}_r^{(v)}$. The fast convergence of the reconstructed signal to the underlying noise-free signal in both cases can be seen. If the initial burn-in period of the first 3000 iterations is ignored, the mean of the residual errors from 3000 to 10000 iterations is 0.7641, with a standard deviation of 0.0537, which implies robust parameter estimation and signal reconstruction. Moreover, the mean value, 0.7641, is related to the SD of the added noise, which is 0.8944.



(a) Actual SCS response in the time domain.

(b) Enlarged actual SCS response in the time domain.

Figure 6.10: *Display of the actual experimental SCS response in the time domain.*

6.6 Estimation of measured ultrasound echo signals

The newly developed estimation system can be applied to measured ultrasound echo signals from both SCSs and MBs. With the combination of the *non-parametric* method and the *parametric* method as illustrated in Figure 6.1 on page 123, the estimated parameters of interest, such as the number of pulse segments and pulse locations, the number of frequencies and their corresponding values for each pulse segment, can be obtained. In this section, one example of measured SCS responses and one example of measured MB responses are shown, and the estimation results are also presented. These two responses were both collected with peak negative pressure of 550kPa and with an incident frequency of 1.83MHz.

6.6.1 Estimation of an SCS response

Figure 6.10(a) displays a typical SCS response, which has a single pulse segment. This typical response is first modeled as a single pulse signal, which is composed of several sinusoids. The proposed estimation system is then carried out and repeated 100 times to the single SCS response. Finally, these results are averaged and calculated to give the means and the SD values of the estimated parameters of interest, i.e. pulse locations, number of frequencies and the corresponding frequency values.

The estimated pulse location in the time domain is presented in Table 6.6 with the means and the SD values. Moreover, the estimates are also marked in Figure 6.10(a) using a red dotted line to denote the start and end points of the pulse segment. Fig-

Estimates(change-points)	1^{st}	2^{nd}
Means	927	1007
SD values	1	4

Table 6.6: *Estimated pulse location for the SCS response*

K	$0 \sim 7$	8	9	10	11	$\geqslant 12$
MAP	0	1	61	35	3	0

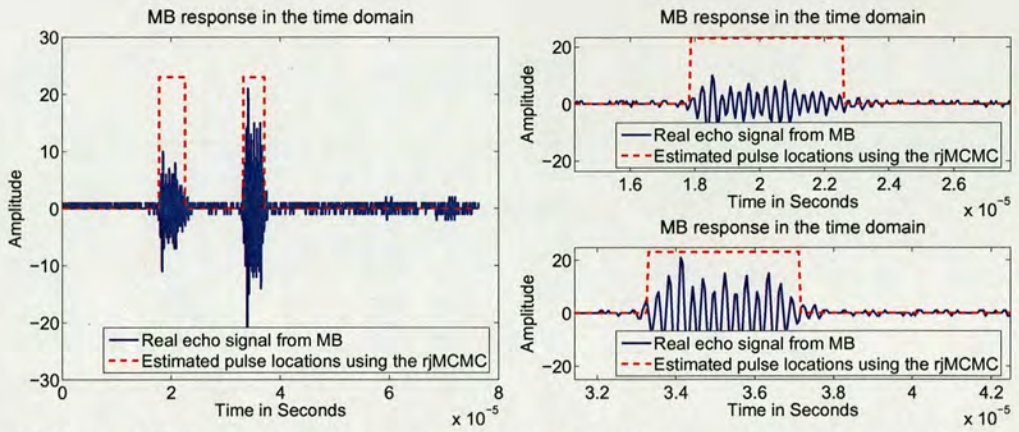
Table 6.7: *Spread of MAP estimates of number of frequency for the SCS response*

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	1.8138	0.0734
2^{nd}	2.1687	0.2858
3^{rd}	2.4339	0.3511
4^{th}	2.6924	0.3516
5^{th}	2.9891	0.4241
6^{th}	3.2320	0.3157
7^{th}	3.3145	0.2756
8^{th}	3.3941	0.4564
9^{th}	3.5345	0.6786

Table 6.8: *Estimated frequency values for the SCS response*

ure 6.10(b) shows an enlarged version of Figure 6.10(a) with explicit indication of the estimated pulse location.

Furthermore, for each run of the estimation algorithm, the standard Bayesian estimate of the model order for a particular pulse segment i, i.e. the number of frequency, is the MAP value of K_i, which is the value of k_i that maximizes $p(k_i|\mathbf{y}_i)$. In this measured SCS response, there is only one pulse segment. The MAP value of K in the single pulse can be obtained in each run of the algorithm. Then with 100 replications of the algorithm, the spread of MAP values are shown in Table 6.7. It can be seen that the most probable number of frequency is 9, which appears 61 times in the 100 replications. Moreover, the estimated frequency values in the single pulse segment are provided in Table 6.8 with the means and the SD values.



(a) Measured MB response in the time domain. (b) Enlarged measured MB response in the time domain.

Figure 6.11: *Display of the actual experimental MB response in the time domain.*

6.6.2 Estimation of an MB response

Unlike SCS responses, most of MB responses are composed of multiple pulse segments. One typical example with two pulse segments is shown in Figure 6.11(a). It is chosen because it can not only describe the procedure of estimation for multiple pulses clearly, but also will not make the computations too complicated using the *parametric* estimation. Therefore, it is a good compromise between an explicit demonstration of the developed estimation algorithm and the computational complexity that a typical example of ultrasound responses requires.

The developed estimation system is also repeatedly applied to the MB response 100 times. The estimation of the number of pulses, and the corresponding pulse locations are given in Table 6.9. The estimated start and end points of two pulse segments are also marked as red dotted lines in Figure 6.11(a). The enlarged version for each single pulse segment is illustrated in Figure 6.11(b) for clarity. Moreover, the spread of MAP values of the number of frequencies for two pulse segments are shown in Table 6.10. The result in Table 6.10(a) indicates that 63% of the time, the detected number of frequencies is 4 for the first pulse. For the second pulse segment, the most probable number of frequency is 5, which appears 58 times in 100 replications of the estimation. Furthermore, the estimated frequency values of the first pulse are presented in Table 6.11(a) and those of the second pulse are presented in Table 6.11(b). These resulting estimates are given with both means and SD values.

Estimates(change-point)	1^{st}	2^{nd}	3^{rd}	4^{th}
Means	351	444	654	730
SD values	2	4	2	2

Table 6.9: *Estimated pulse location for the MB response*

K_1	MAP
$0 \sim 2$	0
3	10
4	63
5	13
6	9
7	3
8	2
$\geqslant 9$	0

(a) Spread of MAP estimates of number of frequency for the 1^{st} pulse

K_2	MAP
$0 \sim 2$	0
3	6
4	27
5	58
6	6
7	3
$\geqslant 8$	0

(b) Spread of MAP estimates of number of frequency for the 2^{nd} pulse

Table 6.10: *Spread of MAP estimates of number of frequency for two pulse segments in the MB response*

6.6.3 Discussion

The developed estimation algorithm has been applied to the measured SCS response in Section 6.6.1 and the measured MB response in Section 6.6.2 respectively. With respect to the estimated pulse locations in the time domain, the estimates exclude the tails of the pulse segments in the full length signal. For example, compare Figure 3.16(a) on page 56 in Section 3.2.4.2 and Figure 6.10(b) in Section 6.6.1. They display the change-point estimates of the same SCS response using *non-parametric* method and the *parametric* method. In Figure 3.16(a), the estimated pulse location include the tail of the pulse segment, which may be induced by the transducer bandwidth used in actual ultrasound experimental measurements. By contrast, the estimated pulse location shown in Figure 6.10(b) exclude the tail of the pulse segment, which is regarded as noise term when using the *parametric* estimation method. The same situation happens to the estimates of the MB responses.

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	2.4390	0.4282
2^{nd}	2.7667	0.4597
3^{rd}	3.1077	0.5327
4^{th}	3.4059	0.4632

(a) Estimated frequency values for the 1^{st} pulse

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	2.0710	0.4228
2^{nd}	2.6467	0.4207
3^{rd}	2.9594	0.4414
4^{th}	3.3252	0.6160
5^{th}	3.3841	0.6210

(b) Estimated frequency values for the 2^{nd} pulse

Table 6.11: *Estimated frequency values of two pulses in the MB response*

From the point of view of frequency estimation, take the SCS response shown in Figure 3.16(a) on page 56 in Section 3.2.4.2 as an example. The DFT spectrum and the multitaper power spectrum of the single pulse SCS response were displayed in Figure 3.20 on page 65 in Section 3.3.5.2. Figure 3.20(b) has lower side lobes than those in Figure 3.20(a) whereas displays wider peaks in the spectrum. Compared to the results obtained from the *non-parametric* estimation method, the estimation results using *parametric* method exhibits more subharmonics and superharmonics. For the MB responses, the estimated frequency components concentrate around the second harmonic and the third harmonic.

Despite the advantages of this newly developed estimation system in both the time and the frequency domains in a *parametric* way, the limitations of it are twofold. First, in terms of the estimation algorithm itself, the limitation lies in the assumptions made in estimating number of pulses in the signal. Multiple pulses in the echo signal are assumed to be well separated and each of the pulse has a minimum length. Specifically, if two pulses are closely-spaced, they will be regarded as one pulse; if the width of a single pulse is less than a certain value, it will be regarded as noise. Moreover, the proposed model is demonstrated to fit the measured ultrasound echo signals well rather than being justified in a strict sense, which needs more validation in the future

work. Second, in accordance with the requirement of ultrasound imaging techniques, due to the influence of ultrasound transducer in the experimental measurements, the estimated spectral contents may not originate from the true behavior of ultrasound scatterers. Therefore, the signal model needs to take account of the ultrasound transducer characteristics in order to examine the real behavior of the scatterers. This will be studied further in detail in Chapter 7.

6.7 Chapter summary

According to the proposed signal model for measured ultrasound echo signals in Chapter 4, and the approximation Bayesian computation methods investigated in Chapter 5, this chapter has developed a specially designed rjMCMC algorithm to estimate the model parameters via Bayesian inference.

By combining the *non-parametric* estimation methods and the *parametric* estimation methods, an estimation system is created. The advantage of it lies in that it allows an automatic estimation of pulse locations in multiple pulse echo signals, and estimation of frequency components for each pulse segment simultaneously. It provides a new look at the time-frequency analysis based on a signal model within a Bayesian framework. Moreover, the proposed *parametric* model and the developed estimation algorithm are able to produce estimated parameters of interest with improved accuracy. In particular, it exhibits improved frequency resolution compared to Fourier analysis based techniques.

Although it has some limitations that two consecutive pulses or overlapped pulses cannot be discriminated, the developed estimation system reveals more characteristics in both the time and the frequency domains. This can lead to the discrimination of measured ultrasound echo signals from MBs and SCSs in future work.

Chapter 7
Spectral estimation of a modified signal model with transducer characteristics

The investigations in previous chapters are all based on experimentally measured ultrasound echo signals, which are the results of original echoes convolved with the impulse response of the ultrasound transducer. This chapter will dedicate to a modified signal model taking account of the transducer characteristics as model parameters. Moveover, a modified rjMCMC algorithm is developed to estimate the modified model parameters. The obtained estimation results are compared to those with no consideration of the transducer in order to demonstrate the superiority of the modified signal model. After the evaluation, the modified estimation algorithm based on the modified signal model, is applied to measured ultrasound echo signals, which can reveal the true spectral contents of responses from ultrasound scatterers.

7.1 Introduction

In Chapter 4, the proposed signal model does not take account the ultrasound transducer characteristics. Moreover, the estimation algorithm developed in Chapter 6 is used to sample from the posterior distributions of model parameters, which is on the basis of the proposed signal model. However, if the band limit property of the receiver [1] is considered, the proposed signal model and the developed estimation algorithm will be inappropriate.

In order to compensate the transducer influence, one intuitive method is designing an (inverse) filter for the ultrasound receiver. If the (inverse) filter is able to perfectly remove the effects of the receiver and thus recover the true spectra of ultrasound echo signals, the estimation algorithm developed in Chapter 6 can be applied to the inverse filtered signals directly. However, the noise gain introduced by the (inverse) filter

[1] Transducer and receiver refer to the same thing in ultrasound measurements and thus are interchangeable in this thesis.

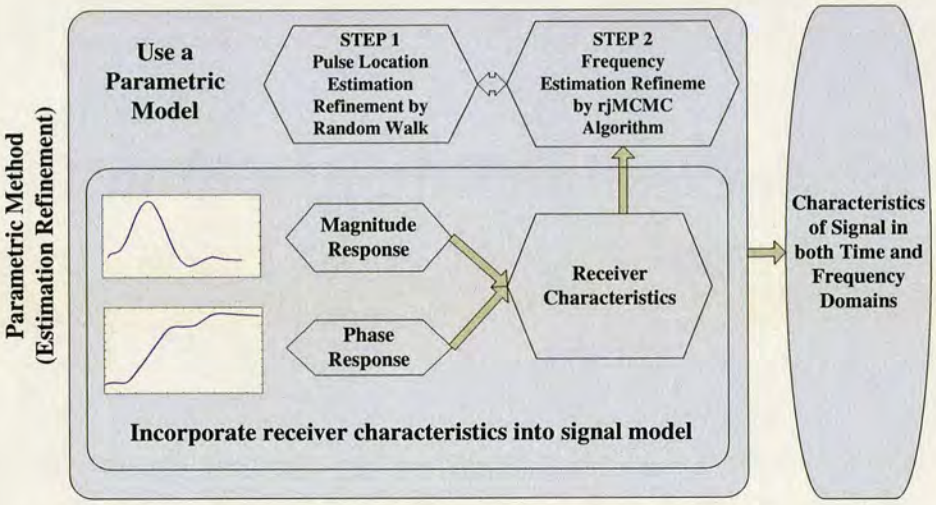


Figure 7.1: *The diagram for the complete estimation for ultrasound echo signals.*

will have an impact on the frequency spectra of ultrasound echo signals. Apart from the noise problem, if using an (inverse) filter, the spectral estimation will consist of two disjoint procedures. First, design an (inverse) filter that is used to recover the original spectra of the measured echo signals; second, apply the estimation algorithm developed in Chapter 6 to extract the spectral content of the recovered signal.

An alternative approach is to incorporate the receiver characteristics into a signal model as model parameters. Then the spectral content of the signal can be estimated using a modified rjMCMC algorithm based on the modified signal model. In this way, the true spectra of ultrasound echo signals will be recovered and the estimation process can be made in a whole single framework. Figure 7.1 illustrates the estimation procedure designed for this modified signal model, which is a part of Figure 1.3 on page 9. This chapter first calibrates the magnitude and the phase responses of the receiver in Section 7.2.1 and Section 7.2.2. Then a brief discussion about the problem associated with the use of an (inverse) filter for the receiver is presented in Section 7.2.3. Having obtained the receiver characteristics, a modified signal model is described in Section 7.3 and the corresponding modified rjMCMC algorithm is investigated to sample from the more complicated posterior distribution in Section 7.4. Finally, the evaluation of the modified estimation algorithm is carried out in Section 7.5.1. This is followed by the application of the estimation algorithm to the measured ultrasound SCS response and MB response in Section 7.5.2.

7.2 Calibration of ultrasound transducer

The calibration of the ultrasound transducer characteristics contains two parts: one is the calibration of the magnitude response and the other is the calibration of the phase response. They are achieved using two different approaches and are investigated in Section 7.2.1 and Section 7.2.2 respectively. The calibration of the ultrasound receiver is based on SCS responses as they are linear scatterers in terms of the transmit frequency.

7.2.1 Magnitude response of the ultrasound receiver

One intuitive method in calibrating the magnitude response of the ultrasound receiver is to calculate the energy density ratio of the measured ultrasound responses and the theoretical responses from SCSs, which will be discussed in detail in Section 7.2.1.1. Unfortunately, using this approach, only finite calibration points are available in the frequency spectrum. Section 7.2.1.2 fits a polynomial curve to these finite values and achieves a continuous curve for the magnitude response, which will be easier to use in the following developed estimation algorithm.

7.2.1.1 Calibration of magnitude response of the ultrasound receiver

A simple illustration of the receiver calibration is depicted in Figure 7.2. The receiver response can be estimated by looking at the signals at the input and output of the transducer. The magnitude response of the receiver can be calculated by taking the ratio of the energy density of the experimentally measured responses from the SCSs [3] and the corresponding theoretical responses [13]. Moreover, since the transmit pulse is a narrow-band signal containing a dominant fundamental and a second harmonic spectral component, the ratio of the magnitude responses will only cover a limited bandwidth in the spectrum. To estimate over the full range of the receiver spectral response, it is necessary to use transmit pulses with different incident frequencies and filter them into the fundamental and the second harmonic spectral components. Therefore, the frequency response of the transducer can be estimated by filtering the transmit pulse into fundamental and second harmonic components, calculating the theoretical echo returns at the transducer input, filtering the measured transducer output in the same way, and then calculating the ratio of the energy density of the output to the input.

To be more specific, for each transmit frequency, energy densities of the filtered fundamental and second harmonic components should be calculated separately. Parseval's theorem indicates that the total energy contained in a waveform, $y(t)$, summed

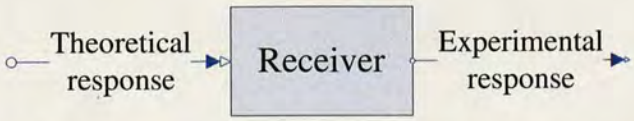


Figure 7.2: *A simple illustration of the receiver calibration.*

across all of time, t, equals to the total energy of the waveform's FT, $Y(f)$, summed across all of its frequency components f, which can be put in the form of:

$$\int_{-\infty}^{+\infty} |y(t)|^2 dt = \int_{-\infty}^{+\infty} |Y(f)|^2 df. \tag{7.1}$$

Concerning a calibration point, $G_{calibrate_{mag}}$, in the frequency spectrum of the ultrasound transducer based on a single frequency, f_{mag}, it can be calculated in the time domain as:

$$G^2_{calibrate_{mag}}(f_{mag}) = \frac{\text{Energy Density}_{exp}}{\text{Energy Density}_{theory}} = \frac{\int_{t_{min}^{(exp)}}^{t_{max}^{(exp)}} |P_{exp}(t)|^2 dt}{\int_{t_{min}^{(theory)}}^{t_{min}^{(theory)}} |P_{theory}(t)|^2 dt}, \tag{7.2}$$

where $P_{exp}(t)$ and $P_{theory}(t)$ represent the measured and the theoretical backscattered radiation in the time domain. The time intervals, i.e. $[t_{min}^{(theory)}, t_{max}^{(theory)}]$ and $[t_{min}^{(exp)}, t_{max}^{(exp)}]$, over which the integration are evaluated should be of sufficient durations to include the complete pulse segment for both the theoretical response and the measured response. A detailed explanation of how to calculate a calibration point in the frequency spectrum of the ultrasound receiver at a specific transmit frequency is presented in Appendix C.

Furthermore, there are 21 different transmit frequencies ranging from 1.2MHz to 4.0MHz. Each of them will be filtered into the fundamental and the second harmonic components and thus there will be 42 calibration points in the spectrum of the ultrasound receiver. Under each transmit frequency, there are several trials of experimental measurements for SCS responses. These SCSs have various radii. With these different trials of measurements, the means and the SD values of the calibration points can be obtained, which are plotted in Figure 7.3. The circles and diamonds denote the means, and the error bars denote the SD values of the calibrated magnitude gains of the receiver. Moreover, the blue circles and error bars represent results from the fundamental components; the red diamonds and error bars represent results from the second har-

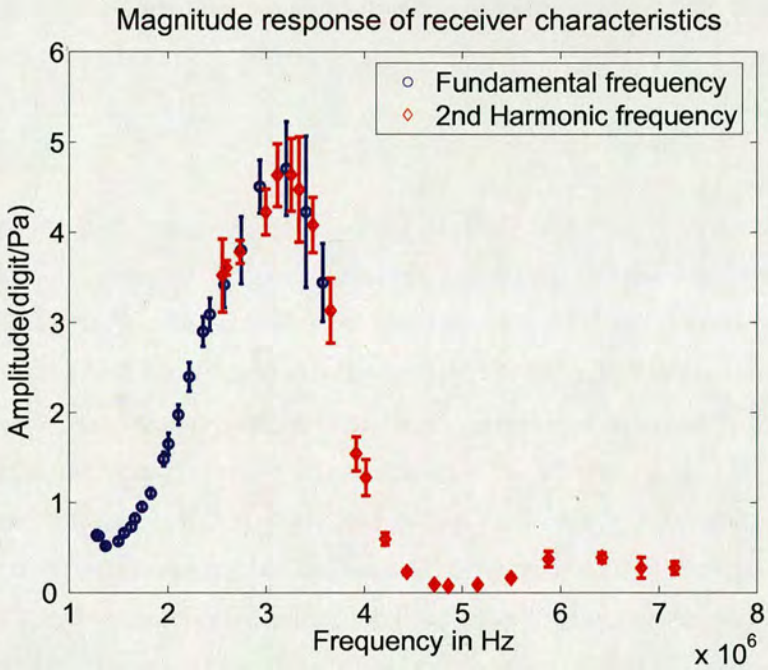


Figure 7.3: *Receiver calibration in the whole frequency range.*

monic components. By checking the overlap of the blue circles and the red diamonds, the accuracy of the experimental measurements can be demonstrated and the shape of the receiver bandwidth can be created. In the full range of the transmit frequencies, the SD values are relatively high when the frequencies are between 2.5MHz and 3.5MHz. This implies that the noise variance in this range will be higher than other frequencies.

7.2.1.2 Fitting a polynomial function to the finite calibration points in the spectrum

In Figure 7.3, there are only finite frequency values rather than a continuous curve available. In order to incorporate this information into a signal model, these finite number of the magnitude gains against frequencies in the spectrum require interpolation. For simplicity, a polynomial function is assumed to fit a curve to the finite numbers in the calibrated magnitude response of the ultrasound receiver. A general expression of the polynomial function can be written as:

$$G(f) = \sum_{p=0}^{P} c_p f^p, \tag{7.3}$$

where c_p, $(p = 0, \ldots, P)$ are the unknown polynomial coefficients, and P is the corresponding model order. This function can also be written in a matrix-vector form and a noise term is also included:

$$\mathbf{y} = \mathbf{Q}\boldsymbol{\beta} + \mathbf{n}, \tag{7.4}$$

where $\boldsymbol{\beta} \triangleq [c_0, c_1, \ldots, c_P]^T$ and $\mathbf{Q} \triangleq [1, f, f^2, \ldots, f^P]$. According to Figure 7.3, the noise is not constant within the whole frequency range. However, as a preliminary curve fitting problem, the noise $\mathbf{n}$ is assumed to be a zero-mean white Gaussian noise with a constant variance of σ_n^2 in order to ease the mathematical calculations.

Traditionally, the most common approach to evaluate the best fit is the least squares (LS) method [51]. It allows for an approximate fit by minimizing the difference between the observations, $\mathbf{y}$, and the fitted curve, $G(f) = \mathbf{Q}\boldsymbol{\beta}$. The coefficients, $\boldsymbol{\beta}$, can be evaluated by solving the standard LS equations. Moreover, there is an alternative to ordinary LS regression, which is called Bayesian linear regression in statistics using Bayesian inference. This inference method within a Bayesian framework has been introduced in Section 4.4.2. In Bayesian linear regression, a prior probability distribution over the signal model parameters, $\boldsymbol{\beta}$, is first employed. Then the posterior distribution of the parameters is calculated based on the prior and the likelihood function. Finally, the estimated model parameters are obtained by being sampled from the posterior density using numerical methods, e.g. Gibbs sampler [108].

According to the proposed signal model in (7.4), the likelihood function can be presented as:

$$p(\mathbf{y}|\boldsymbol{\beta}, \sigma_n^2) = \frac{1}{(2\pi\sigma_n^2)^{N/2}} \exp\left[-\frac{(\mathbf{y} - \mathbf{Q}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{Q}\boldsymbol{\beta})}{2\sigma_n^2}\right]. \tag{7.5}$$

Where N is the number of observations. For ease of the derivation, a uniform distribution and a Jeffrey's prior are assigned to the coefficients, $\boldsymbol{\beta}$, and the noise variance, σ_n^2, respectively. Based on Bayes's rule, the posterior distribution of the coefficients, $\boldsymbol{\beta}$, can be obtained:

$$p(\boldsymbol{\beta}|\mathbf{y}, \sigma_n^2) \propto (2\pi\sigma_n^2)^{-N/2} \exp\left[-\frac{(\mathbf{y} - \mathbf{Q}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{Q}\boldsymbol{\beta})}{2\sigma_n^2}\right] \times \frac{1}{\sigma_n^2}. \tag{7.6}$$

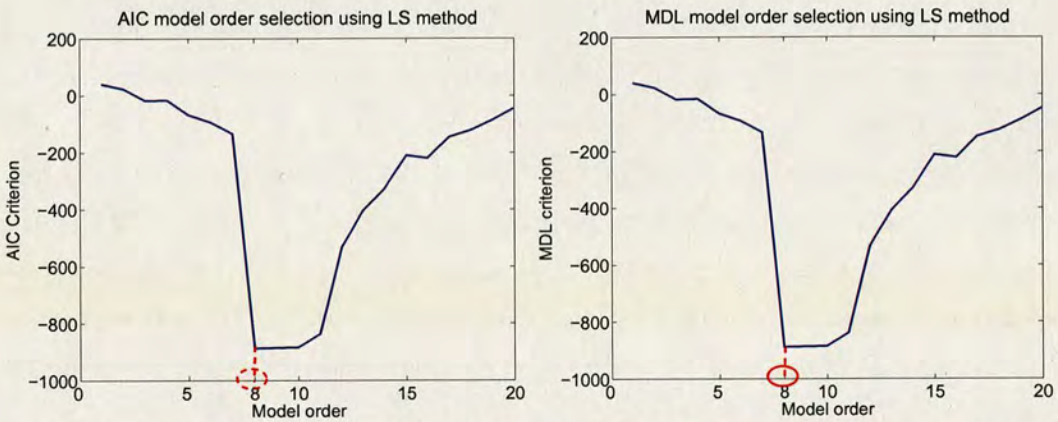
As suggested in [108], in order to generate jointly distributed samples to compute the marginal density of the coefficients, $\boldsymbol{\beta}$, the condensed Gibbs sampler is used. It is a variant of the Gibbs sampler. Using this method, the linear parameters, $\boldsymbol{\beta}$, are

sampled altogether at once. Then the noise variance, σ_n^2, is sampled. Concerning the linear coefficients, β, they can be estimated by drawing a random sample vector from a multivariate Gaussian distribution with the mean, $\mu_\beta = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{y}$, and the covariance matrix, $\Sigma_\beta = \sigma_n^2(\mathbf{Q}^T\mathbf{Q})^{-1}$. Moreover, the estimated noise variance can be sampled from a conditional density, $\sigma_n^2|\beta, \mathbf{y} \sim \mathcal{IG}\left(\frac{N}{2}, \frac{1}{2}[\mathbf{y} - \mathbf{Q}\beta]^T[\mathbf{y} - \mathbf{Q}\beta]\right)$. All these derivations using Bayesian linear regression can be found in [108]. It should be noted that, since the prior of the coefficient vector is uniformly distributed in this case, the maximisation of the posterior distribution is equivalent to the estimation using LS method.

Furthermore, when fitting a suitable curve to the finite numbers in the spectrum, whether the estimation approach is based on the LS method or based on Bayesian linear regression, determining a model order for the polynomial function is required in the first place. In Section 4.6.1, some simple and commonly used model order selection criteria have been examined briefly. Among these criteria, AIC and MDL are two popular ones and have been demonstrated with good performances in solving simple problems [90]. Figure 7.4 illustrate the model order selection for the magnitude response of the receiver using both AIC and MDL criteria. The lowest point at $P = 8$ in either AIC criterion or MDL criterion indicates it is the best choice for the model order for polynomial coefficients in the proposed model in (7.4). The estimated results of the polynomial coefficients are displayed in Table 7.1 with means and SD values. The means of the estimated coefficients are used to create a continuous curve for the magnitude response of the receiver, as displayed in Figure 7.5. It shows a good agreement between the finite number of blue circles and the red continuous line. Unfortunately, the interpolation introduces negative values for the magnitude response. They are forced to be zeros in the following estimation procedure. This will not have significant influence on the resulting estimated frequency components.

7.2.2 Phase response of the ultrasound receiver

By calculating the ratio of energy density of the measured ultrasound echo signals to the theoretical echo signals, only the magnitude response can be obtained. However, the magnitude response of the receiver cannot be used to fully characterise the ultrasound receiver. Estimation of the phase response of the receiver is also necessary. There is a different way to estimate the phase response, which is calibrated by taking the ratio of the FT of the measured responses to the theoretical responses. It is often



(a) Model order selection for magnitude response using AIC criterion.

(b) Model order selection for magnitude response using MDL criterion.

Figure 7.4: *Model order selection for the magnitude response of the receiver.*

Coefficients $\hat{\beta}$	Means	SD values
$\hat{c}_0$	-180.6035	9.3645
$\hat{c}_1$	505.7294	24.4812
$\hat{c}_2$	-578.6615	26.5858
$\hat{c}_3$	352.5759	15.6980
$\hat{c}_4$	-124.9146	5.5281
$\hat{c}_5$	26.5174	1.1927
$\hat{c}_6$	-3.3228	0.1545
$\hat{c}_7$	0.2266	0.0110
$\hat{c}_8$	-0.0065	3.3221×10^{-4}

Table 7.1: *Estimated polynomial coefficients of the magnitude response of the receiver.*

termed as a transfer function approach. As the FT of a signal covers the whole range of frequencies, from 0MHz to $fs/2 \approx 9.8$MHz, the transmit pulse which is used to generate the theoretical response, and the measured response do not need to be filtered to the fundamental and the second harmonic separately. The procedure of estimating the phase response based on the receiver transfer function (RTF) using a single measured response and a single theoretical response from SCSs can be summarised as:

1. Calculate the phase response of an experimentally measured echo signal based on the FT.

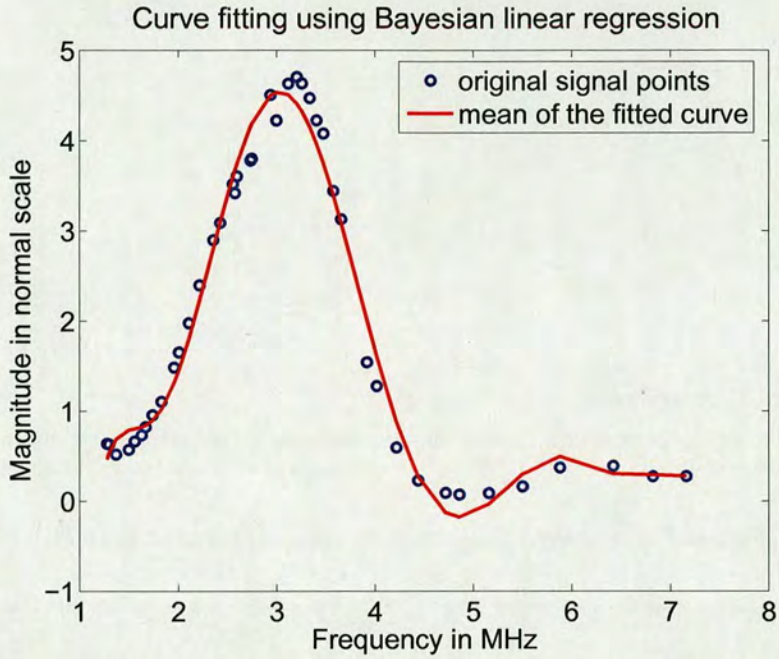


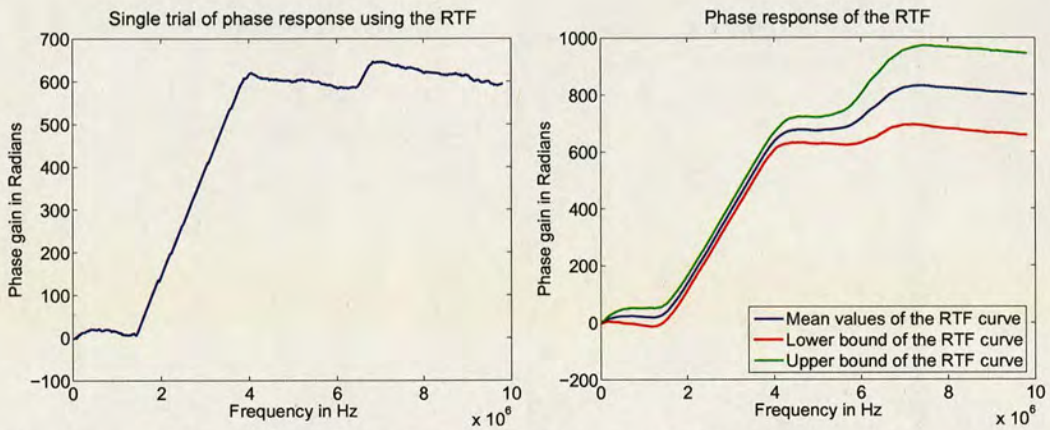
Figure 7.5: *Curve fitting for individual experimental data with known model order.*

2. Calculate the phase response of a theoretical echo signal based on the FT.

3. Calculate the the phase response RTF by calculating the difference between the phase response of the output and that of the input to the ultrasound receiver according to (7.7):

$$G_{calibrate_{phase}} = \arg[H(\omega)] = \arg[F_{output}(\omega)] - \arg[F_{input}(\omega)]. \tag{7.7}$$

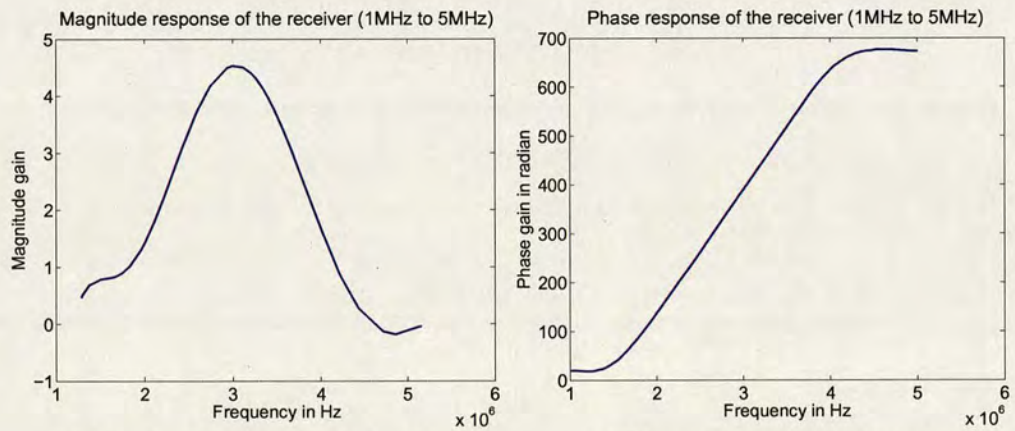
As aforementioned, under each transmit frequency, there are several trials of measured SCSs responses with various radii and a single theoretical SCS response available. For example, there are 9 measured SCS responses with a transmit frequency at 1.67MHz. Therefore, there will be 9 phase responses, which can produce a mean RTF of the phase response. Moreover, considering 21 different transmit frequencies, there will be 21 RTFs of the phase response. Average these phase responses to produce a overall mean RTF of the phase response and the corresponding SDs. Figure 7.6(a) displays an example with the obtained phase response under the transmit frequency at 1.67MHz. Figure 7.6(b) depicts the overall averaged RTF of the obtained phase responses and the corresponding SDs. All these displayed phase responses are unwrapped so as to



(a) Averaged phase response of the receiver obtained with one transmit frequency at 1.67MHz.

(b) Averaged phase response of the receiver obtained with 21 different transmit frequencies.

Figure 7.6: *Averaged phase response obtained based on the RTF*



(a) Magnitude response of the receiver.

(b) Phase response of the receiver.

Figure 7.7: *Magnitude and phase responses of the receiver from 1MHz to 5MHz.*

correct phase angles to produce smoother phase plots.

7.2.3 Why not use an inverse filter?

As the region of interest (ROI) in characterising the spectrum of the ultrasound receiver is from 1.0MHz to 5.0MHz [3], the estimated magnitude and phase responses of the receiver with this ROI are shown in Figure 7.7. In order to remove the receiver influence on the responses from ultrasound scatterers, the most intuitive approach is to design an (inverse) filter for the ultrasound receiver.

Although there are many techniques available to tackle the problem of designing

an optimal (inverse) filter, some restrictions and limitations are still inevitable. For example, only when a system, denoted as $H(Z)$ in the Z-plane, [2] is a minimum phase system, which requires that all the zeros of $H(Z)$ should be constrained to be within the unit circle of the Z-plane. However, it is not an easy job to justify whether the ultrasound receiver used in the measurements is a stable system or not. Furthermore, (inverse) filtering is a noise enhancement process. If a white noise, $w(n)$, with a variance of σ_n^2 is passed through a filter, $h(n)$, then the output can be given by the convolution $y(n) = \sum h(\tau)w(n-\tau)$. Therefore,

$$\mathbf{E}[y^2(n)] = \sum_{\tau}\sum_{\hat{\tau}} h(\tau)h(\hat{\tau})\mathbf{E}[w(n-\tau)w(n-\hat{\tau})] \tag{7.8}$$

and, since $\mathbf{E}[w(n-\tau)w(n-\hat{\tau})] = \sigma_n^2\delta(\hat{\tau}-\tau)$, the noise gain is given by:

$$P_{NG} = \sum h^2(\tau). \tag{7.9}$$

As a result, designing an (inverse) filter is not an optimal approach. From the perspective of improving the signal model and making it more suitable for the measured ultrasound responses, instead of removing the influence of the ultrasound receiver, the magnitude and the phase responses can be included in the signal model design as part of the model parameters. In this way, there is no need to justify whether the receiver is a stable system and the noise gain problem will be circumvented as well. A modified signal model with transducer characteristics and the modified rjMCMC algorithms for estimating the model parameters will be discussed in detail in Section 7.3 and Section 7.4.

7.3 A modified signal model with transducer characteristics

A new signal model taking account of the ultrasound transducer characteristics is proposed, which modifies the signal model expressed in (4.17a) on page 91. The new model does not require designing an (inverse) filter, and it combines the process of removing the transducer effect and estimating the spectral content of a signal jointly in a single framework.

[2] In mathematics, Z-plane is a geometric representation of the complex numbers composed of the real part and the orthogonal imaginary part.

7.3.1 Proposed signal model and its likelihood function

In this section, a modified signal model which incorporates the transducer characteristics is proposed. For ease of the problem, only a single pulse segment is considered and it is easy to extend a single pulse signal to a full length signal model with multiple pulse segments.

7.3.1.1 Modified signal model

Consider a single pulse segment in the signal model in (4.19) on page 92 , a modified signal model incorporating the transducer characteristics can be given as:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{H}\mathbf{s} + \mathbf{n} \\ &= \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k + \mathbf{n}, \end{aligned} \tag{7.10}$$

where $\boldsymbol{D}$ and $\boldsymbol{a}_k$ have the same definitions as $\boldsymbol{D}_i$ and $\boldsymbol{a}_{k_i}$ in (4.19) on page 92 in Section 4.5.2. $\boldsymbol{D}$ contains frequency information ω_k, and $\boldsymbol{a}_k$ contains the amplitudes for each frequency. Moreover, $\mathbf{n}$ is a zero-mean additive white Gaussian noise. The only difference between (4.19) and (7.10) is the receiver matrix $\mathbf{H}$, which contains the impulse response of the ultrasound receiver in the time domain. The matrix $\mathbf{H}$ is a convolution matrix, [3] which is defined as:

$$\mathbf{H} \triangleq \begin{bmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \dots & \vdots \\ \vdots & h(1) & \dots & 0 \\ h(M-1) & \vdots & \dots & h(0) \\ \vdots & h(M-1) & \dots & h(1) \\ 0 & \vdots & \dots & \vdots \\ 0 & 0 & \dots & h(M-1) \end{bmatrix}, \tag{7.11}$$

where

$$\mathbf{h} = [h(0), \quad h(1), \quad \dots, \quad h(M-1)]^T \tag{7.12}$$

is the impulse response of the calibrated ultrasound receiver in the time domain. It can be easily achieved by being transformed from the frequency spectrum of the receiver $\tilde{\mathbf{H}}$

[3] A convolution matrix is a matrix formed from a vector, which is able to convert the convolution operation of two vectors to the multiplication operation of two in the time domain.

in the frequency domain. As the magnitude and the phase responses can be assigned polynomial models as a function of frequency bins, $\tilde{\mathbf{H}}$ in the frequency domain can be obtained with the coefficient vectors $\boldsymbol{\varphi}_m$ and $\boldsymbol{\varphi}_p$:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{H}}(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p) = \left(\sum_{r=0}^{P_1} \boldsymbol{\varphi}_m(r) \boldsymbol{\omega}^r \right) \cdot \exp \left[j \left(\sum_{s=0}^{P_2} \boldsymbol{\varphi}_p(s) \boldsymbol{\omega}^s \right) \right], \tag{7.13}$$

where $\boldsymbol{\omega} \triangleq [\omega(0), \quad \omega(1), \quad \ldots, \quad \omega(M-1)]^T$ represents a vector of M frequency bins. $\boldsymbol{\varphi}_m$ and $\boldsymbol{\varphi}_p$ are coefficients of the polynomial functions for the magnitude response and the phase response of the receiver respectively. They are assumed to have model order P_1 and P_2.

7.3.1.2 Likelihood function

It is assumed that the noise vector $\mathbf{n}$ is an i.i.d. white Gaussian noise with a variance of σ_k^2, and the parameters describing the received echo signal are stationary throughout the entire observation. The likelihood function of the signal model in (7.10) can be expressed as:

$$p(\mathbf{y}_i | k, \boldsymbol{\psi}_k) = (2\pi\sigma_k^2)^{-L/2} \cdot \exp \left\{ -\frac{1}{2\sigma_k^2} [\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k]^T [\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k] \right\}, \tag{7.14}$$

where $L = M + N - 1$. N is the length of the original unfiltered echo signal and M is the length of the receiver in the time domain. k is the number of sinusoids in the signal. Other parameters are denoted as: $\boldsymbol{\psi}_k \triangleq (\boldsymbol{\theta}_k, \boldsymbol{\varphi}_r)$, in which $\boldsymbol{\theta}_k \triangleq (\boldsymbol{\omega}_k, \boldsymbol{a}_k, \sigma_k^2)$ is defined the same as $\boldsymbol{\theta}_{k_i}$ in (4.20) on page 93, and $\boldsymbol{\varphi}_r \triangleq (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p)$. The initial guesses for the coefficients, $\boldsymbol{\varphi}_r$, are known from the calibration of the ultrasound transducer.

7.3.2 Prior distributions for model parameters

A hierarchical structure for the parameters of the modified signal model, is illustrated in Figure 7.8 using a DAG. v and γ are assigned fixed values; Λ and δ^2 are random variables. They have the same definitions as v_i, γ_i, Λ_i and δ_i^2, that are displayed in Figure 6.2 on page 124. $\boldsymbol{\varphi}_m$ and $\boldsymbol{\varphi}_p$ are new parameters compared to Figure 6.2. Since both the magnitude and the phase responses are assumed to be in the form of polynomial functions, $\boldsymbol{\varphi}_m$ and $\boldsymbol{\varphi}_p$ represent their polynomial coefficients respectively, which are independent with each other. According to Figure 7.8, the joint prior distribution

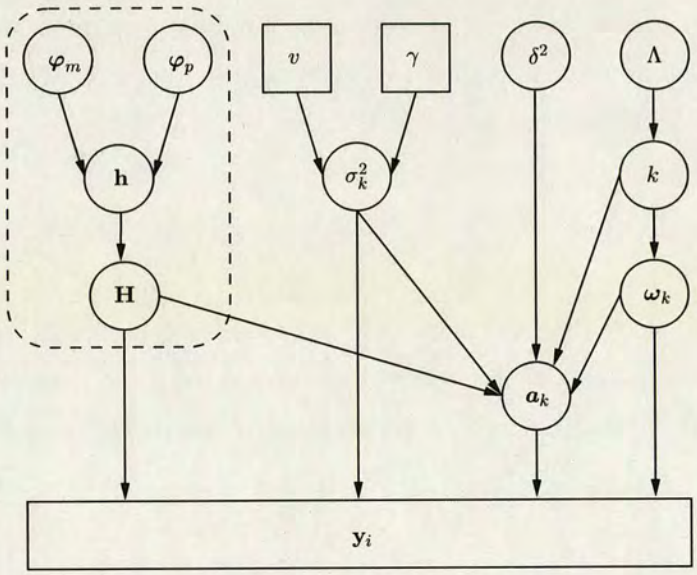


Figure 7.8: *DAG illustrating the hierarchical structure for model parameters of the modified signal model.*

of all these parameters can be expressed as:

$$p(k, \boldsymbol{\psi}_k) = p(k, \boldsymbol{\theta}_k, \boldsymbol{\varphi}_r) = p(k, \boldsymbol{\theta}_k|\boldsymbol{\varphi}_r)p(\boldsymbol{\varphi}_r), \tag{7.15}$$

where $p(k, \boldsymbol{\theta}_k|\boldsymbol{\varphi}_r)$ is given by (6.4a) on page 125 if $0 < k \leqslant k_{max}$. If $k = 0$, $p(k, \boldsymbol{\theta}_k|\boldsymbol{\varphi}_r)$ is given by (6.4b). There are some rough ideas about polynomial coefficients of the magnitude and the phase responses in Section 7.2.1 and Section 7.2.2, in which case, a Gaussian distribution centered about a single value is a good choice of the prior for a single particular coefficient parameter.

In order to simplify the expression of the joint prior distribution and the joint posterior distribution in the following expressions, $p(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p)$ will remain in the intact form.

7.3.3 Posterior distributions for model parameters

Based on Bayes's rule, by multiplying (7.14) and (7.15), the joint posterior distribution can be obtained as:

$$p(k, \boldsymbol{\psi}_k|\mathbf{y}_i) \propto (2\pi\sigma_k^2)^{-L/2} \cdot \exp\left\{-\frac{1}{2\sigma_k^2}[\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k]^T[\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k]\right\}$$
$$\times\, p(k, \boldsymbol{\theta}_k|\boldsymbol{\varphi}_r)p(\boldsymbol{\varphi}_r). \tag{7.16}$$

The derivation and simplification of this joint posterior distribution are similar to those in Section 6.2.2, which are described in detail in Appendix D. After the simplification, the joint posterior distribution after integration is expressed as:

$$\begin{aligned} p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p | \mathbf{y}_i, 0 < k \leqslant k_{max}) \propto & (\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i)^{-(1+L+v)/2} \\ & \times \frac{(\Lambda/((\delta^2+1)\pi))^k}{k!} \times \Lambda^{\epsilon_1 - 1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \\ & \times \delta^{-2(\alpha_{\delta^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta^2}}{\delta^2}\right] \\ & \times p(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p), \end{aligned} \tag{7.17}$$

where

$$\boldsymbol{P}_k^{(r)} = \boldsymbol{I}_{L\times L} - \mathbf{H}\boldsymbol{D}\boldsymbol{M}_k^{(r)}\boldsymbol{D}^T\mathbf{H}^T \tag{7.18a}$$

$$\left\{\boldsymbol{M}_k^{(r)}\right\}^{-1} = (1+\delta^{-2})\boldsymbol{D}^T\mathbf{H}^T\mathbf{H}\boldsymbol{D} \tag{7.18b}$$

$$\boldsymbol{m}_k^{(r)} = \boldsymbol{M}_k^{(r)}\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i, \tag{7.18c}$$

and $\boldsymbol{P}_0^{(r)} \triangleq \boldsymbol{I}_{L\times L}$, in which L is the length of the signal after being convolved with the receiver $\mathbf{H}$. In the case of $k = 0$, $\boldsymbol{a}_0^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{a}_0 \triangleq 0$ and $|2\pi\sigma_k^2\boldsymbol{\Sigma}_0|^{1/2} \triangleq 1$ are adopted in computation of the posterior distribution, as suggested in [33, 34]. Therefore, the posterior distribution is presented as:

$$\begin{aligned} p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p | \mathbf{y}_i, k = 0) \propto & \frac{1}{\sigma_k^{1+L}} \cdot \exp\left(-\frac{\mathbf{y}_i^T\mathbf{y}_i + \gamma}{2\sigma_k^2}\right) \cdot \sigma_k^{-2(v/2+1)} \\ & \times \Lambda^{\epsilon_1 - 1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \\ & \times p(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p). \end{aligned} \tag{7.19}$$

7.4 Spectral estimation by sampling from the posterior distributions

A rjMCMC algorithm introduced in [122] is used to perform Bayesian computation in jointly detecting the desired model order and estimating the model parameters by sampling from their posterior distributions. In Chapter 6, how to apply the rjMCMC algorithm to spectral estimation via sampling from the posterior distribution of frequencies has already been investigated. In this chapter, although the signal model is

modified, the spectral estimation process is similar while it requires sampling from a more complicated posterior distribution in (7.17).

7.4.1 Estimation of a frequency component using a modified rjMCMC algorithm

In order to sample from the complicated posterior distribution, which takes account the receiver characteristics, a similar procedure using a rjMCMC algorithm can be carried out. The modified rjMCMC algorithm splits the sampling of model parameters into two main parts: a dimension-changing part and a fixed-dimension part. In the former part, the algorithm samples the number of frequencies, k, from $p(k, \boldsymbol{\omega}_k|\mathbf{y}_i, \Lambda, \delta^2, \boldsymbol{\varphi}_r)$; in the latter part, it samples the frequency values $\boldsymbol{\omega}_k$ from $p(\boldsymbol{\omega}_k|\mathbf{y}_i, k, \Lambda, \delta^2, \boldsymbol{\varphi}_r)$. The procedure of the rjMCMC algorithm for spectral estimation of a sum of sinusoids signal model has been described in Section 6.2.3. This section concentrates on the difference between the modified rjMCMC algorithms for the modified signal model, and the original developed rjMCMC algorithm. Moreover, the update move for the ultrasound receiver parameters, $\boldsymbol{\varphi}_r$, which will be examined in detail in Section 7.4.2.

7.4.1.1 Update move at a constant model order

In this *update* move for a frequency component, the model order, k, is fixed. According to the MH algorithm, the acceptance ratio of this move is expressed as:

$$\begin{aligned} r_{update} &\triangleq \text{(posterior distribution ratio)} \times \text{(proposal ratio)} \\ &= \frac{p(k, \boldsymbol{\omega}^*_{j,k}|\mathbf{y}_i, \boldsymbol{\varphi}_r)}{p(k, \boldsymbol{\omega}_{j,k}|\mathbf{y}_i, \boldsymbol{\varphi}_r)} \times \frac{q(\omega_{j,k}|\omega^*_{j,k})}{q(\omega^*_{j,k}|\omega_{j,k})}. \end{aligned} \tag{7.20}$$

Compared to (6.11) on page 130, the same proposal ratio is adopted. A hybrid MH step is used: $\lambda q_1(\cdot) + (1-\lambda)q_2(\cdot)$, in which $q_1(\cdot)$ represents a global update based on the frequency spectrum of the observations, $\mathbf{y}$, and $q_2(\cdot)$ represents a local update using a random walk algorithm. They can be calculated using (6.10a) and (6.10b) on page 129.

The difference lies in the posterior distribution ratio. Since the full conditional posterior distribution can be stated as: $p(k, \boldsymbol{\omega}_k|\boldsymbol{\varphi}_r, \mathbf{y}_i) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\varphi}_r|\mathbf{y}_i)$, and according to (7.17), the acceptance ratio can be expanded as:

$$r_{update} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \{\boldsymbol{P}_k^{(r)}\}^* \mathbf{y}_i} \right)^{(1+L+v)/2} \times \frac{q(\omega_{j,k}|\omega^*_{j,k})}{q(\omega^*_{j,k}|\omega_{j,k})}. \tag{7.21}$$

Algorithm 7.1: Update Move for a frequency

1 Propose a candidate frequency $\omega_{j,k}^*$ and the new frequency vector becomes: $\boldsymbol{\omega}_k^* = [\boldsymbol{\omega}_{1:j-1}^{(i)}, \omega_{j,k}^*, \boldsymbol{\omega}_{j+1,k}^{(i)}]$;
2 Evaluate α_{update} in terms of (7.21) and (7.22) ;
3 Sample $u \sim U_{[0,1]}$;
4 **if** $u \leqslant \alpha_{update}$ **then**
5 | The state of the Markov chain at iteration $i+1$ becomes $(k, \boldsymbol{\omega}_k^*)$;
6 **else**
7 | It remains at $(k, \boldsymbol{\omega}_k^{(i)})$;
8 **end**

With the acceptance ratio of the *update* move, its probability of accepting this move is defined as:

$$\alpha_{update} = \min\{1, r_{update}\}. \tag{7.22}$$

The schema for an *update* move for a frequency component is summarized in Algorithm 7.1.

7.4.1.2 Dimension-changing moves

The same as in Section 6.2.3.2, the acceptance ratio of the dimension-changing move is written as:

$$r_{change} \triangleq \text{(posterior distribution ratio)} \times \text{(proposal ratio)} \times \text{(Jacobian)}. \tag{7.23}$$

In this equation, the proposal ratio has the same definition as in Section 6.2.3.2 and can be calculated using (6.21) on page 132. The Jacobian term is one because the dimension only increases and decreases by one in each iteration. Therefore, the difference in dimension-changing moves lies in the posterior distribution ratio.

Since $p(k, \boldsymbol{\omega}_k | \Lambda, \delta^2, \boldsymbol{\varphi}_r, \mathbf{y}_i) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\varphi}_r | \mathbf{y}_i)$, and using (7.17), the posterior distribution ratio can be obtained as:

$$\frac{p(k+1, \boldsymbol{\omega}_{k+1} | \Lambda, \delta^2, \boldsymbol{\varphi}_r, \mathbf{y}_i)}{p(k, \boldsymbol{\omega}_k | \Lambda, \delta^2, \boldsymbol{\varphi}_r, \mathbf{y}_i)} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \boldsymbol{P}_{k+1}^{(r)} \mathbf{y}_i} \right)^{1+L+v} \times \frac{\Lambda}{(k+1)(\delta^2+1)\pi}. \tag{7.24}$$

Accordingly, the acceptance ratio of the dimension-changing moves can be achieved

Algorithm 7.2: Birth Move for a frequency

1 Propose a candidate frequency ω^* on $(0, \pi)$: $\boldsymbol{\omega}_{k+1} = [\boldsymbol{\omega}_k, \omega^*]$;
2 Evaluate α_{birth} according to (7.25) and (7.26a) ;
3 Sample $u \sim U_{[0,1]}$;
4 **if** $u \leqslant \alpha_{birth}$ **then**
5 | The state of the Markov chain at iteration $i+1$ becomes $(k+1, \boldsymbol{\omega}_{k+1})$;
6 **else**
7 | It remains at $(k, \boldsymbol{\omega}_k)$;
8 **end**

Algorithm 7.3: Death Move for a frequency

1 Remove the jth frequency component from the existing frequencies randomly. The candidate frequency vector becomes: $\boldsymbol{\omega}_{k+1}^{(i+1)} = [\boldsymbol{\omega}_{(1:j-1,k+1)}^{(i)}, \boldsymbol{\omega}_{(j+1:k+1,k+1)}^{(i)}]$;
2 Evaluate α_{death} according to (7.25) and (7.26b) ;
3 Sample $u \sim U_{[0,1]}$;
4 **if** $(u \leqslant \alpha_{death})$ **then**
5 | The state of the Markov chain at iteration $i+1$ becomes $(k, \boldsymbol{\omega}_k)$
6 **else**
7 | It remains at $(k+1, \boldsymbol{\omega}_{k+1})$;
8 **end**

by substituting (6.21) and (7.24) into (7.23):

$$r_{change} = \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \boldsymbol{P}_{k+1}^{(r)} \mathbf{y}_i} \right)^{1+L+v} \times \frac{1}{(k+1)(\delta^2+1)}. \tag{7.25}$$

Furthermore, the probabilities of accepting a *birth* move and a *death* move can be presented as follows:

$$\alpha_{birth} = \min\{1, r_{change}\} \tag{7.26a}$$

$$\alpha_{death} = \min\{1, r_{change}^{-1}\}. \tag{7.26b}$$

The schemas for a *birth* move and a *death* move are summarized in Algorithm 7.2 and Algorithm 7.3 respectively.

7.4.2 Update move for ultrasound receiver parameters φ_r

In this section, $0 < k \leqslant k_{max}$ is assumed. If $k = 0$, the receiver parameters are sampled from their prior distribution $p(\varphi_r) = p(\varphi_m, \varphi_p)$. Since

$$p(\varphi_m, \varphi_p | k, \omega_k, \Lambda, \delta^2, \mathbf{y}_i) \propto p(k, \omega_k, \Lambda, \delta^2, \varphi_m, \varphi_p | \mathbf{y}_i). \tag{7.27}$$

and via (7.17), the posterior distribution from which the samples can be drawn is given by:

$$\begin{aligned} p(\varphi_m, \varphi_p | k, \omega_k, \Lambda, \delta^2, \mathbf{y}_i) \propto & (\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i)^{-(1+L+v)/2} \cdot \frac{\Lambda^k}{k! \pi^k} \cdot \left(\frac{1}{1+\delta^2} \right)^k \\ & \times \Lambda^{\epsilon_1 - 1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \times \delta^{-2(\alpha_{\delta^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta^2}}{\delta^2} \right] \\ & \times p(\varphi_m, \varphi_p), \end{aligned} \tag{7.28}$$

where $\boldsymbol{P}_k^{(r)}$ is the matrix stated in (7.18a) and evaluated with the parameter values (k, ω_k, Λ, δ^2, φ_r).

As the receiver parameters, φ_m and φ_p have been obtained with rough estimates in Section 7.2.1 and Section 7.2.2, the priors for φ_m and φ_p are both assigned Gaussian distributions centered on the rough estimates. Moreover, they are assumed to be independent with each other: $p(\varphi_m, \varphi_p) = p(\varphi_m)p(\varphi_p)$. Also, single elements in φ_m and φ_p are independent with one another, such that:

$$p(\varphi_m) = p(\varphi_{m_1}) p(\varphi_{m_2}) \cdots p(\varphi_{m_{P_1}}) \tag{7.29a}$$

$$p(\varphi_p) = p(\varphi_{p_1}) p(\varphi_{p_2}) \cdots p(\varphi_{p_{P_2}}). \tag{7.29b}$$

Therefore, the prior for each single element in φ_m and φ_p can be expressed as:

$$\varphi_{m_i} \sim \mathcal{N}(\varphi_{m_i}^{(init)}, \sigma_m^2) \tag{7.30a}$$

$$\varphi_{p_i} \sim \mathcal{N}(\varphi_{p_i}^{(init)}, \sigma_p^2), \tag{7.30b}$$

where i denotes the index in vector φ_m and φ_p. $\varphi_{m_i}^{(init)}$ and $\varphi_{p_i}^{(init)}$ are the polynomial coefficients of the magnitude and the phase responses obtained in Section 7.2.1 and Section 7.2.2. They are treated as the initial values in updating the ith element in φ_m and φ_p respectively in the modified rjMCMC algorithm. σ_m^2 and σ_p^2 relate to the

uncertainties in the estimates of the initial values of $\varphi_{m_i}^{(init)}$ and $\varphi_{p_i}^{(init)}$.

Each element φ_{m_i} in $\boldsymbol{\varphi}_m$ or φ_{p_i} in $\boldsymbol{\varphi}_p$ does not depend on other elements, and thereby the *update* move for a single element can be applied to all elements in the same way. The same as the *update* move for a frequency, which is described in Section 7.4.1.1, the acceptance ratio for the updating can be expressed as:

$$\begin{aligned} r_\varphi(\varphi_i, \varphi_i^*) &\triangleq \text{posterior distribution ratio} \times \text{proposal ratio} \\ &= \frac{p(\varphi_i^*|k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y}_i)}{p(\varphi_i|k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y}_i)} \times \frac{q_\varphi(\varphi_i|\varphi_i^*)}{q_\varphi(\varphi_i^*|\varphi_i)} \\ &= \left(\frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i}{\gamma + \mathbf{y}_i^T \{\boldsymbol{P}_k^{(r)}\}^* \mathbf{y}_i} \right)^{(1+L+v)/2} \times \frac{p(\varphi_i^*)}{p(\varphi_i)} \times \frac{q_\varphi(\varphi_i|\varphi_i^*)}{q_\varphi(\varphi_i^*|\varphi_i)}, \end{aligned} \tag{7.31}$$

where $\boldsymbol{P}_k^{(r)}$ and $\{\boldsymbol{P}_k^{(r)}\}^*$ are matrices with different receiver parameters, $(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p)$ and $(\boldsymbol{\varphi}_m^*, \boldsymbol{\varphi}_p^*)$, respectively, which are calculated using (7.18a). Moreover, $q_\varphi(\cdot)$ is the proposal distribution for a single element in the receiver parameters, φ_i. For ease of the computation, it is chosen as a random walk on the space of the polynomial coefficients in the magnitude response or the phase response. In other words, each proposed coefficient depends on its previous value, the update is expressed as:

$$\varphi_i^*|\varphi_i \sim \mathcal{N}(\varphi_i, \sigma_\varphi^2), \tag{7.32}$$

where σ_φ^2 is a constant that determines the step size in updating φ_i. Therefore, the acceptance probability for updating a single element, φ_i, is given by:

$$\alpha_\varphi(\varphi_i, \varphi_i^*) = \min\{1, r_\varphi(\varphi_i, \varphi_i^*)\}. \tag{7.33}$$

The schema for an *update* move for a single coefficient φ_i is summarized in Algorithm 7.4.

7.4.3 Update move for other related parameters

The rest of the parameters in the modified signal model are nuisance parameters $\boldsymbol{a}_k$ and σ_k^2, and hyperparameters Λ and δ^2. They have the same definitions as those discussed in Section 6.2.3.3 and can be derived in the same way. In general, there is no need to update the nuisance parameters $\boldsymbol{a}_k$ and σ_k^2 as they have already been integrated out from the joint posterior distribution. Unfortunately, they are required

Algorithm 7.4: Update Move for a coefficient of the receiver polynomial function

Propose a candidate frequency φ_i^* according to (7.32) ;
Evaluate $\alpha_\varphi(\varphi_i, \varphi_i^*)$ in terms of (7.31) and (7.33) ;
Sample $u \sim U_{[0,1]}$;
if $u \leqslant \alpha_\varphi(\varphi_i, \varphi_i^*)$ **then**
    The state of the Markov chain at iteration $i+1$ becomes φ_i^* ;
else
    It remains at φ_i ;
end

Parameters	Prior Distribution	Posterior Distribution
$\boldsymbol{a}_k$	$\mathcal{N}(0, \delta^2 \boldsymbol{D}^T \mathbf{H}^T \mathbf{H} \boldsymbol{D})$	$\mathcal{N}(\boldsymbol{m}_k^{(r)}, \sigma_k^2 \boldsymbol{M}_k^{(r)})$
σ_k^2	$\mathcal{IG}(v/2, \gamma/2)$	$\mathcal{IG}(\frac{1+L+v}{2}, \frac{\gamma + \mathbf{y}_i^T \boldsymbol{P}_k^{(r)} \mathbf{y}_i}{2})$
Λ	$\mathcal{G}a(1/2+\epsilon_1, \epsilon_2)$	$\mathcal{G}a(1/2+k+\epsilon_1, 1+\epsilon_2)$
δ^2	$\mathcal{IG}(\alpha_{\delta^2}, \beta_{\delta^2})$	$\mathcal{IG}(k+\alpha_{\delta^2}, \frac{1}{2\sigma_k^2}(\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k)^T(\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k) + \beta_{\delta^2})$

Table 7.2: *Update the parameters*

when updating hyperparameters Λ and δ^2. Moreover, the priors for them are chosen as conjugate priors so as to make the computation much easier. Their priors and the corresponding posteriors are stated in Table 7.2. The derivation of $\boldsymbol{a}_k$ and σ_k^2 can be found in Appendix D. The analysis of updating Λ and δ^2 are the same as those addressed in Section 6.2.3.3, and thus will not be repeated again.

7.4.4 Procedure of spectral estimation based on the modified signal model

In summary, with the incorporation of the receiver characteristics as the signal model parameters, the procedure of spectral estimation for the modified signal model using the modified rjMCMC algorithm is stated in Algorithm 7.5.

7.5 Estimation results

In this section, a simulated signal is created to evaluate the performance of this modified rjMCMC algorithm based on the modified signal model, which incorporates the receiver characteristics. Then the modified estimation algorithm is applied to experimentally measured ultrasound echo signals, including both SCS responses and MB responses.

Algorithm 7.5: Spectral estimation for the modified signal model using a rjM-CMC algorithm

Initialization: set $(\{k, \psi_k\}^{(0)})$;
Iteration
for $i = 1$ **to** *numIteration* **do**
  Sample $u \sim U_{[0,1]}$;
  if $u \leqslant b_{k^{(i)}}$ **then**
    Perform *birth* move of a new frequency according to Algorithm 7.2 on page 174 ;
  else if $(u \leqslant b_{k^{(i)}} + d_{k^{(i)}})$ **then**
    Perform *death* move of an existing frequency according to Algorithm 7.3 on page 174 ;
  else
    Perform *update* move of each frequency component according to Algorithm 7.1 on page 173 ;
  end
  Update move of the receiver parameters φ_m and φ_p according to (7.31) and (7.33), and Algorithm 7.4;
  Sample related parameters $\boldsymbol{a}_k$, σ_k^2, Λ and δ^2 ;
end

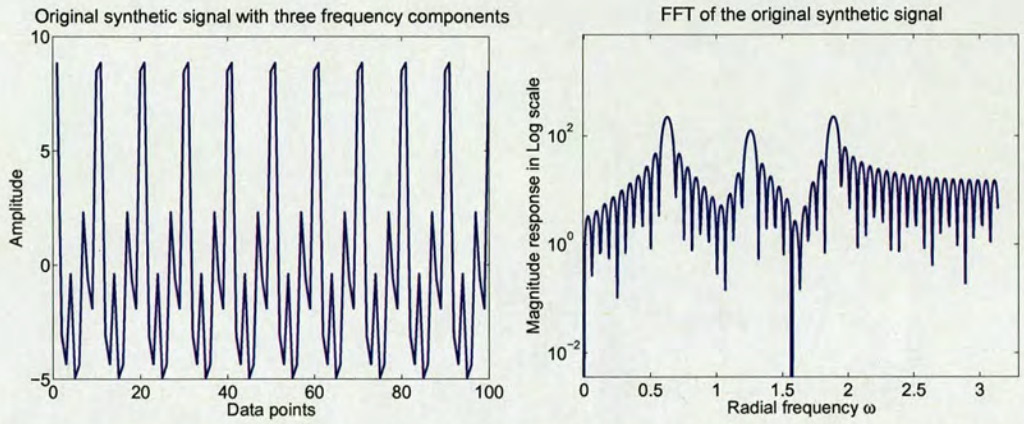
7.5.1 Evaluation of the estimation algorithm on a simulated signal

When a simulated signal is used, the ground truths of the model parameters are known. Therefore, the estimated frequencies can be compared to their true values to examine how well the algorithm performed. Moreover, the estimates obtained from the modified estimation algorithm are also compared to the results achieved from the direct estimation algorithm in Chapter 6.

7.5.1.1 Estimation results for a simulated signal

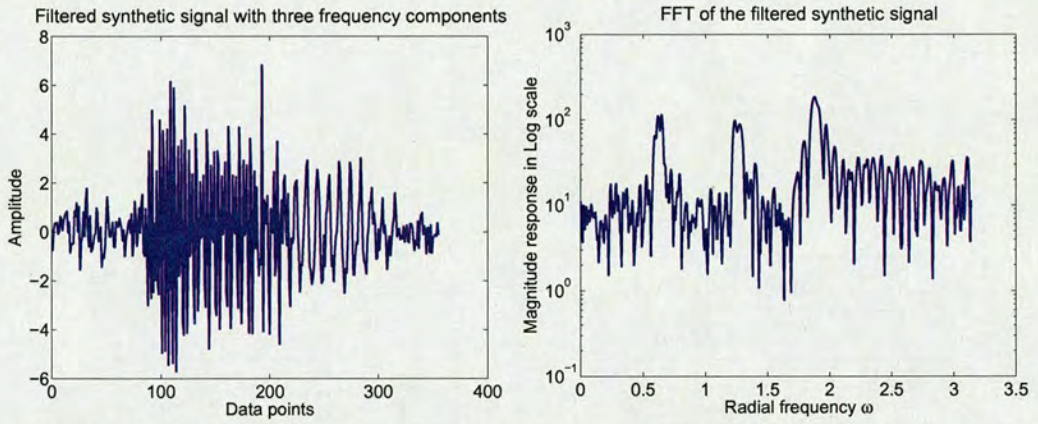
The simulated signal has a length of $N = 100$ data points, and has three frequency components: 0.2π, 0.4π and 0.6π. Its parameter setting is displayed in Table 7.3. The original simulated signal and its frequency spectrum are shown in Figure 7.9(a) and Figure 7.9(b) respectively. Then the original synthetic signal convolves with the impulse response of the receiver, which is shown in Figure 7.10. The obtained filtered synthetic signal after being convolved with the calibrated ultrasound receiver in the time domain is shown in Figure 7.9(c) and the SNR is 20dB. Its corresponding frequency spectrum is also shown in Figure 7.9(d).

Furthermore, as the calibrated receiver has a length of $M = 256$ data points, the



(a) The original simulated noisy free signal in the time domain.

(b) The original simulated noisy free signal in the frequency domain.



(c) Filtered simulated signal in the time domain.

(d) Filtered simulated signal in the frequency domain.

Figure 7.9: *The original noise free signal and the filtered signal with the calibrated ultrasound receiver in both the time and the frequency domain.*

k	$E_k = a_{s,k}^2 + a_{c,k}^2$	$-\arg\tan(a_{s,k}/a_{c,k})$	ω_k
1	20	0	0.2π
2	6.3246	$\pi/4$	0.4π
3	20	$\pi/3$	0.6π

Table 7.3: *Parameter settings for the simulated signal.*

filtered synthetic signal has a length of $L = N + M - 1 = 355$ data points. The parameter setup of this algorithm has the same values as those shown in Table 6.2 on page 141, except for the receiver parameters, φ_m and φ_p. Since the magnitude and the phase responses are modeled as polynomial functions, the initial estimated values for

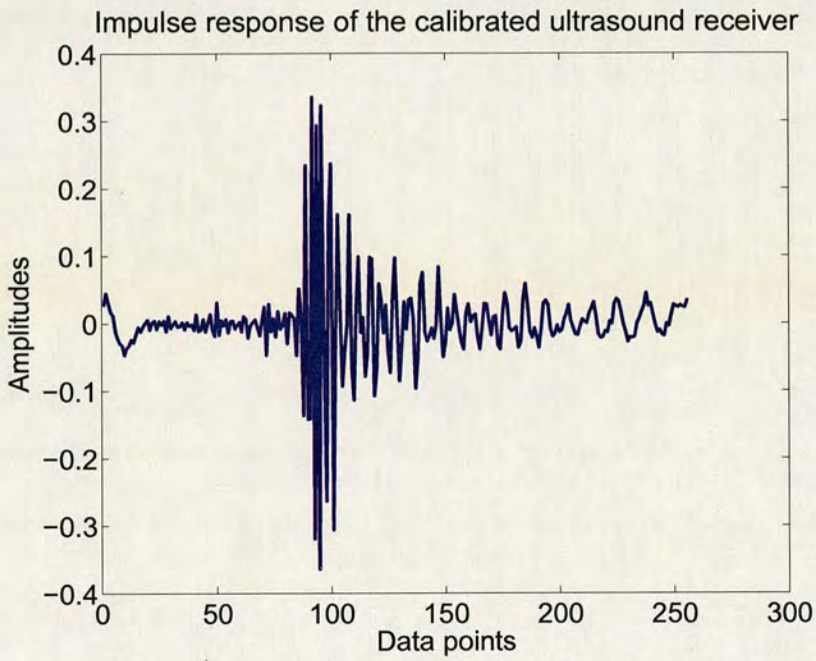


Figure 7.10: *Impulse response of the calibrated ultrasound receiver.*

$\varphi_m^{(init)}$	Coefficient values
$\hat{\varphi}_{m_1}$	-180.6035
$\hat{\varphi}_{m_2}$	505.7294
$\hat{\varphi}_{m_3}$	-578.6615
$\hat{\varphi}_{m_4}$	352.5759
$\hat{\varphi}_{m_5}$	-124.9146
$\hat{\varphi}_{m_6}$	26.5174
$\hat{\varphi}_{m_7}$	-3.3228
$\hat{\varphi}_{m_8}$	0.2266
$\hat{\varphi}_{m_9}$	-0.0065

(a) Initial coefficient values for the magnitude response of the receiver

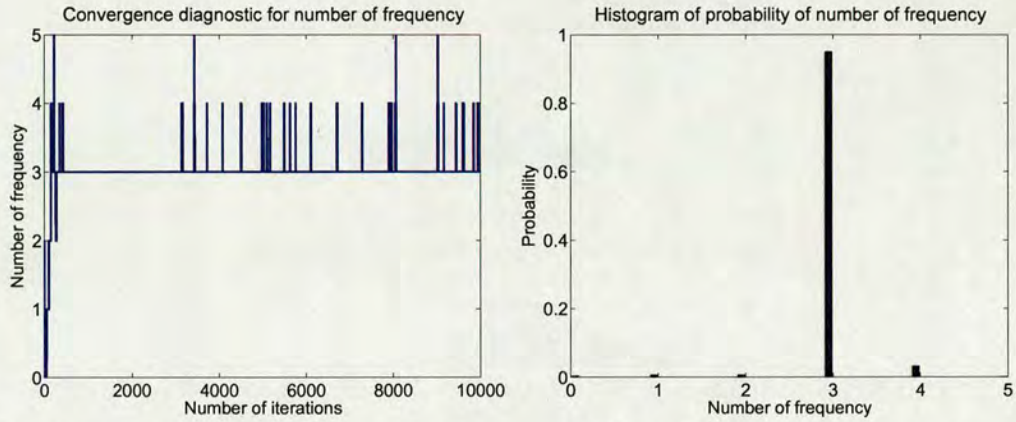
$\varphi_p^{(init)}$	Coefficient values
$\hat{\varphi}_{p_1}$	$-1.984e+3$
$\hat{\varphi}_{p_2}$	0.0078
$\hat{\varphi}_{p_3}$	$-1.2204e-8$
$\hat{\varphi}_{p_4}$	$9.9765e-15$
$\hat{\varphi}_{p_5}$	$-4.6898e-21$
$\hat{\varphi}_{p_6}$	$1.3206e-27$
$\hat{\varphi}_{p_7}$	$-2.1938e-34$
$\hat{\varphi}_{p_8}$	$1.9747e-41$
$\hat{\varphi}_{p_9}$	$-7.3994e-49$

(b) Initial coefficient values for the phase response of the receiver

Table 7.4: *Initial values for the coefficients of the magnitude and the phase responses of the receiver*

their coefficients $\varphi_m^{(init)}$ and $\varphi_p^{(init)}$ are given in Table 7.4, which have been achieved in Section 7.2.

Consider the simulated signal with a single noise realisation, following the pro-



(a) Convergence diagnostic for number of frequency.

(b) Histogram of number of frequency.

Figure 7.11: *Estimation of number of frequency based on a modified signal model with one noise realisation.*

cedure described in Algorithm 7.5, a Markov chain, is constructed. After a certain number of iterations, samples from the desired posterior distribution can be drawn. Then the MAP estimator is used to estimate the model parameters, especially the parameters of interests, i.e. the number of frequencies k and their corresponding values ω_k. Seen from Figure 7.11(a), the constructed Markov chain converged quickly and the burn-in period is around 1000 iterations. By discarding the first 1000 samples from its equilibrium distribution, the remaining 9000 samples are used to plot a histogram of the number of frequency in the original signal. Figure 7.11(b) displays this histogram, indicating the estimated number of frequency in the original unfiltered signal is three, with the highest probability. Moreover, Figure 7.12 displays histograms of the estimated frequency values. The peak values in the histograms indicate the most probable frequency values in the original signal.

Furthermore, there are 100 different noise realisations used to check the stability of the modified rjMCMC algorithm. Table 7.5 shows the spread of MAP estimates of the number of frequency for the original signal using 100 Monte Carlo runs. It can be inferred that 91% of the time, the detected number of frequency is 3. Moreover, the estimated frequency values are compared to the ground truths, which are shown in Table 7.6. The estimated means and SD values are also presented. The errors between the means and the ground truths are small, and the small SD values indicate good stability of the modified estimation algorithm.

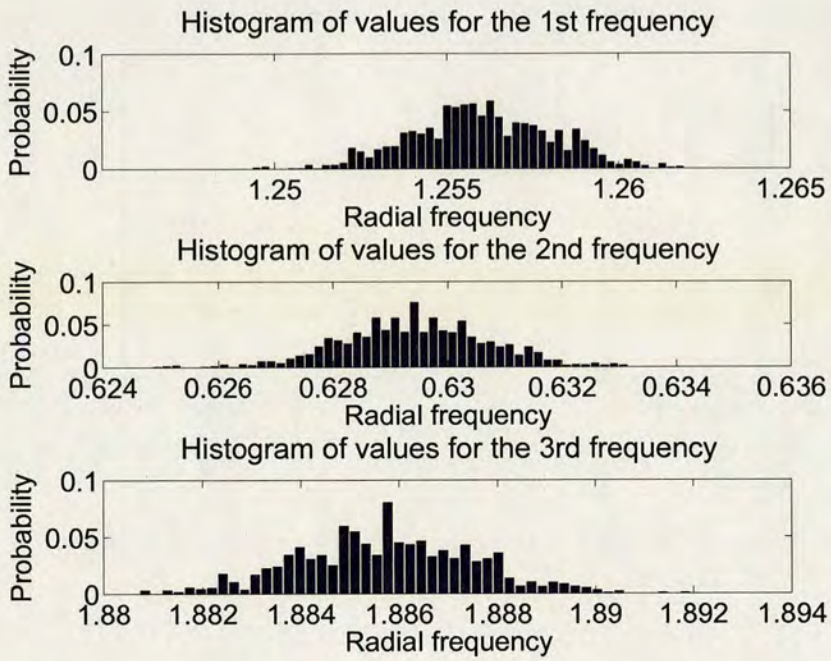


Figure 7.12: *Histograms of frequency values based on a modified signal model with one noise realisation.*

K	$0 \sim 2$	3	4	5	$\geqslant 6$
MAP	0	91	8	1	0

Table 7.5: *Spread of MAP estimates of number of frequency for the original signal with* 100 *noise realisations*

7.5.1.2 Comparison of estimation results between the direct estimation and the modified estimation with various SNRs

In Section 7.5.1.1, the simulated filtered signal taken as the input of the estimation algorithm is the result of a convolution of original synthetic signal and the impulse response of the ultrasound receiver. The direct estimation algorithm proposed in Chapter 6 is applied to the truncated filtered signal. The duration of truncation is determined by the estimated pulse location given by the rjMCMC algorithm. The original unfiltered synthetic signal has a finite length, which is equivalent to being multiplied by a rectangular window. Due to this inevitable window effect, when the unfiltered signal is convolved with the impulse response of the transducer, there will be a small tail in the filtered signal although it is not obvious. This tail comes from the cut-off effect in the unfiltered signal being convolved with the transducer impulse response, as

Frequency components	Ground Truth	Estimated values (mean $\pm$ SD)
1^{st} Freq.	$0.2\pi = 0.6283$	0.6323 ± 0.0099
2^{nd} Freq.	$0.4\pi = 1.2566$	1.2451 ± 0.0877
3^{rd} Freq.	$0.6\pi = 1.8850$	1.8723 ± 0.0895

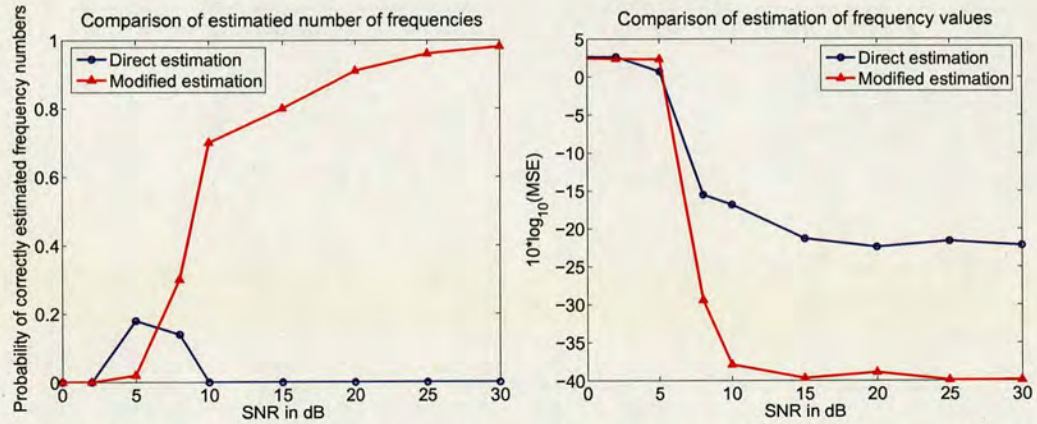
Table 7.6: *Comparison of frequency values between the ground truth and the estimates*

K	$0 \sim 3$	4	5	6	7	8	$\geqslant 9$
MAP	0	28	34	24	12	2	0

Table 7.7: *Spread of MAP estimates of number of frequency with* 100 *noise realisations using the direct estimation algorithm.*

shown in Figure 7.10, and thus may have certain influence on the estimated frequency components. Therefore, to make the comparison in a more reasonable way, the proposed rjMCMC estimation algorithm is applied to the filtered signal in Figure 7.9(c) for both temporal and spectral contents. In this way, the spectral contents are estimated using the truncated filtered signal.

When the SNR is 20dB, the rjMCMC algorithm is applied 100 times on the simulated signal with 100 different noise realisations. The spread of MAP estimates of the number of frequency using the direct estimation proposed in Chapter 6 is displayed in Table 7.7. The most probable number of frequencies is 5 as it has the highest probability amongst others. However, the ground truth in the unfiltered signal is 3. Furthermore, the comparison is also made with various SNRs from 0dB to 30dB. For each SNR, with 100 noise realisations, the probabilities of correctly estimating the number of frequencies against SNRs are shown in Figure 7.13(a). They are denoted as a red line marked by triangles. The probabilities of correctly estimating the number of frequencies using the direct estimation algorithm developed in Chapter 6 are represented using a blue line marked by circles. The modified rjMCMC estimation algorithm based on the modified signal model has a better performance against various SNRs. With the increase of the SNR value, the probability of correct detection of frequency numbers also increases. When the SNR is at 10dB, the correct detection of frequency numbers is 70 out of the total 100 replications of the estimation algorithm. Moreover, when the SNR is 30dB, the correctly detected number of frequencies is 98%, indicating good accuracy



(a) Probability of correct estimation of number of frequency for different SNR.

(b) MSE of the estimated frequency values for different SNR.

Figure 7.13: *Compare the estimated number of frequencies and their values between the direct estimation and the modified estimation.*

and stability of the modified estimation algorithm.

In addition to the number of frequencies, the estimated frequency values can also be obtained simultaneously against various SNRs. For each SNR, the error is calculated between the estimated frequency values and the ground truths. They are computed by taking the logarithm of the MSE values, which are shown in Figure 7.13(b). The red line marked by triangles represents the logarithm of MSE using the modified estimation algorithm based on the modified signal model. It is also compared to a blue line marked by circles, which illustrates the logarithm of MSE using the direct estimation with no consideration of receiver characteristics. It can be inferred that when the SNR is below 5dB, the performance of the correct estimation of frequency values of both the direct estimation and the modified estimation are poor. When the SNR is higher than 5dB, both the performance of direct estimation and the modified estimation algorithms increase. Moreover, the accuracy of the modified algorithm has an improvement of 18dB.

7.5.2 Estimation of measured ultrasound echo signals

The evaluation of the modified rjMCMC estimation algorithm based on the modified signal model has been addressed in Section 7.5.1. It was demonstrated that if a signal had been convolved with the impulse response of the calibrated ultrasound receiver, the modified estimation algorithm using the modified signal model is superior to the

direct estimation algorithm with no consideration of the receiver. This is because the modified model accounts for the phase variation introduced by the transducer. With this modified model and its corresponding modified rjMCMC estimation algorithm, the true spectral contents of measured ultrasound echo signals can be recovered and extracted correctly. Therefore, the real behavior of ultrasound scatterers will be revealed.

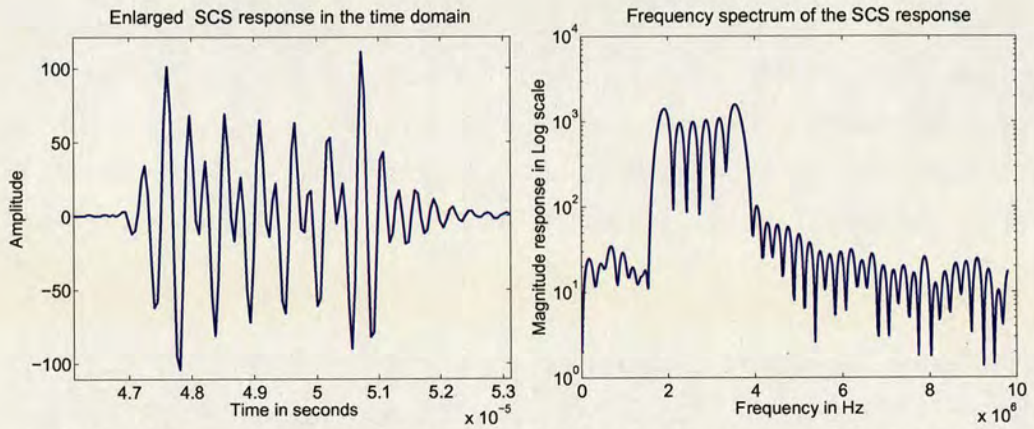
7.5.2.1 Estimation of the SCS responses

In the frequency spectrum of measured ultrasound echo signals from SCSs, due to the bandwidth limit of the ultrasound receiver, the frequencies outside the frequency range, i.e. from 1MHz to 5MHz, will be attenuated and thus might be regarded as side lobes rather than the main lobe peaks. Take one SCS response with a transmit frequency, which has a fundamental frequency at 1.83MHz as an example. Figure 7.14(a) displays an enlarged version of the SCS response in the time domain. The developed spectral estimation algorithm is repeatedly applied to this SCS response 100 times. The spread of the MAP estimates of the number of frequency is given in Table 7.8. Moreover, the estimated frequency values are displayed in Table 7.9. Their means and SD values can be computed based on the results from 100 replications of the modified rjMCMC estimation algorithm.

Compared to the frequency spectrum of this SCS response in Figure 7.14(b), one frequency component is detected by the modified estimation algorithm, 3.8425MHz. However, it is neglected by the estimation without consideration of the transducer due to the attenuation region in the frequency spectrum. Moreover, the obvious four peaks between 2.0MHz and 4.0MHz in the spectrum are suppressed when the receiver bandwidth is considered. Only two of them are regarded as main lobes and the other two peaks are neglected as side lobe noise using the modified rjMCMC estimation algorithm. Although the resulting estimates cannot be justified as there is no ground truths for the frequency components in the echo signals, it is demonstrated that the transducer does have certain impact on the frequency estimation.

7.5.2.2 Estimation of the MB responses

With respect to measured ultrasound echo signals from MBs, there will be a misunderstanding of MB behavior due to the incorrect spectral estimation. Therefore, the modified rjMCMC estimation algorithm based on the modified signal model is trying to



(a) Enlarged SCS response in the time domain. (b) FFT of SCS response in the frequency domain.

Figure 7.14: *Display of the measured SCS response in the time and the frequency domains.*

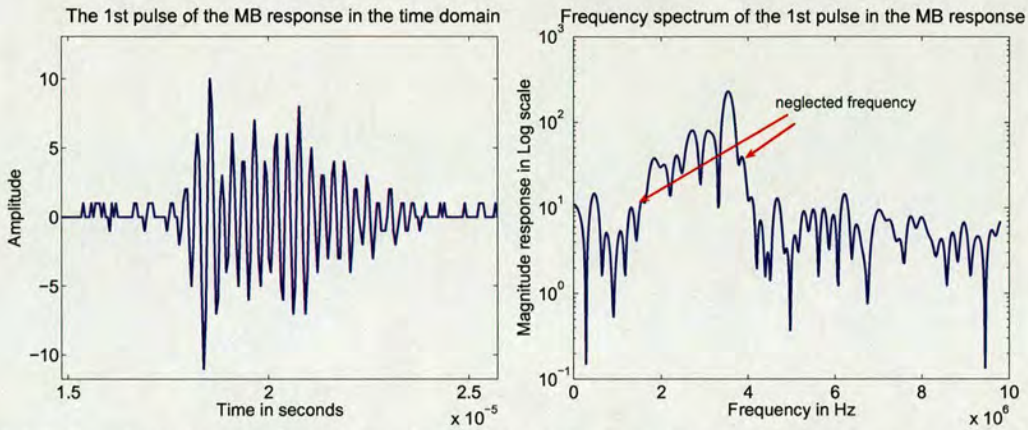
K	$0 \sim 1$	2	3	4	5	$\geqslant 6$
MAP	0	4	10	40	46	0

Table 7.8: *Spread of MAP estimates of number of frequency in the SCS response with* 100 *repeats of the algorithm.*

recover the original spectra of MB responses from the distortion and finally reveal the true MB behavior from experimental measurements. Take one MB response with an incident pulse, which has a fundamental frequency at 1.83MHz as an example. There are two pulse segments in the MB response. They are displayed in Figure 7.15(a) and Figure 7.16(a) respectively. As the modified signal model and the modified estimation algorithm are only valid for a single pulse duration, these two pulse segments in the MB response are separated manually and then are estimated individually. The same

Freq. components	Means (in MHz)	SD values (in MHz)
1st Freq.	1.6772	0.3876
2nd Freq.	2.3684	0.5427
3rd Freq.	3.0139	0.5653
4th Freq.	3.4766	0.1926
5th Freq.	3.8425	0.7956

Table 7.9: *Estimated frequency values of the measured SCS response*



(a) Enlarged version of the 1^{st} pulse of the MB response in the time domain.

(b) FFT of the 1^{st} pulse of the MB response in the frequency domain.

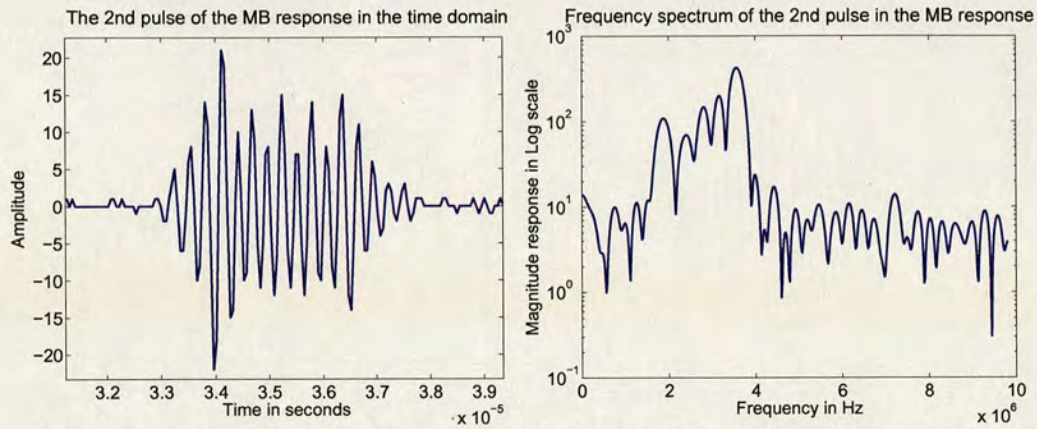
Figure 7.15: *Display of the 1^{st} pulse of the measured MB response in the time and the frequency domains.*

as the estimation performed for the SCS response, the algorithm is carried out 100 times on two pulses in the MB response. Table 7.10 displays the spread of the MAP estimates of the number of frequency for two MB pulse segments. The results indicate there are 7 frequency components in the first pulse and 5 in the second pulse as they have the highest probabilities. Furthermore, the means of the estimated frequency values for two pulses are presented in Table 7.11, as well as their corresponding SD values. These two MB pulse segments have similar spectral contents although their signal shape in the time domain are different. Moreover, they have more harmonics than those in the SCS response, as predicted in ultrasonics literature.

7.5.2.3 Discussion about the results

The estimation results obtained using the modified rjMCMC algorithm based on the modified signal model have been compared to the results achieved in Chapter 6. The proposed signal model and the developed estimation algorithm examined in Chapter 6 do not take account the ultrasound receiver characteristics. However, the receiver has an influence on the frequency spectra of measured ultrasound echo signals. The estimation results shown in this chapter are able to recover the frequency components attenuated by the receiver.

Compare the estimated frequency values of the SCS response in Table 6.8 on page 152 and Table 7.9. The frequency, 3.8425MHz, was detected in Table 7.9 whereas was ne-



(a) Enlarged version of the 2^{nd} pulse of the MB response in the time domain.

(b) FFT of the 2^{nd} pulse of the MB response in the frequency domain.

Figure 7.16: *Display of the 2^{nd} pulse of the measured MB response in the time and the frequency domains.*

K_1	MAP
$0 \sim 3$	0
4	2
5	8
6	28
7	60
8	2
$\geqslant 9$	0

(a) Spread of MAP estimates of number of frequency for the 1^{st} pulse

K_2	MAP
$0 \sim 1$	0
2	2
3	2
4	32
5	38
6	26
$\geqslant 7$	0

(b) Spread of MAP estimates of number of frequency for the 2^{nd} pulse

Table 7.10: *Spread of MAP estimates of number of frequency for two pulse segments in the MB response*

glected in Table 6.8. It is because this frequency component locates in the frequency attenuation range caused by the ultrasound receiver, which can be seen in the frequency spectrum of the SCS response in Figure 7.14(b).

Moreover, with respect to the MB response, two pulse segments have similar spectral contents although more frequencies are detected in the first pulse. Compare the estimated frequencies of the first pulse shown in Table 6.11(a) on page 155 and Table 7.11(a). The detected smallest frequency component, 1.6869MHz, and the largest

Freq. components	Means (in MHz)	SD values (in MHz)
1^{st}	1.6869	0.5187
2^{nd}	2.0637	0.3797
3^{rd}	2.5152	0.4463
4^{th}	2.9093	0.3228
5^{th}	3.1695	0.2719
6^{th}	3.4006	0.2398
7^{th}	3.8942	0.2456

(a) Estimated frequency values for the 1^{st} pulse

Freq. component	Means (in MHz)	SD values (in MHz)
1^{st}	1.8478	0.0623
2^{nd}	2.5615	0.3464
3^{rd}	2.9011	0.2867
4^{th}	3.3241	0.2791
5^{th}	3.6147	0.1601

(b) Estimated frequency values for the 2^{nd} pulse

Table 7.11: *Estimated frequency values for two pulse segments in the MB response*

one, 3.8942MHz, have been detected when using the modified algorithm based on the modified signal model. They are both neglected in Table 6.11(a), however, because the smallest and the largest frequency values are suppressed a lot by the ultrasound receiver. In addition to the estimates in the first pulse segment, the spectral content estimated in the second pulse segment of the MB response display similar estimated frequency values in Table 6.11(b) and Table 7.11(b). The detected smallest frequency, 1.8478MHz and the largest one, 3.6147MHz in Table 7.11(b) are closer to the fundamental and the second harmonic frequencies, compared to those estimates presented in Table 6.11(b).

As a result, it is demonstrated that the ultrasound receiver does have a certain impact on the spectral content of measured ultrasound echo signals, no matter they are from SCSs or MBs.

7.6 Chapter summary

This chapter has proposed a modified signal model that takes the receiver characteristics as model parameters, and has developed a modified rjMCMC estimation algorithm to extract the spectral contents in measured ultrasound echo signals. The modified model and the modified estimation algorithm are shown to be robust in estimating the spectral content of ultrasound scatterers. Moreover, they have been compared to the proposed model and the developed corresponding estimation algorithm in Chapter 6, which have demonstrated their superiority. Some frequency components, which are suppressed by the limited bandwidth of the ultrasound receiver, can be recovered. Although this chapter only considers a single pulse segment in the ultrasound responses, it is easy for the signal model and the modified estimation algorithm to extend to include all pulse segments in multiple pulse echo signals.

From the perspective of ultrasound contrast imaging, according to the analysis of the spectral contents of measured ultrasound echo signals, the real behavior of scatterers, especially the MBs, have been revealed . This is able to broaden the research area of ultrasound MBs. Furthermore, it will also assist in discrimination of echo signals from MBs and SCSs.

Chapter 8
Conclusions and future work

This thesis has been concerned with ultilising statistical signal processing techniques to characterise experimentally measured echo signals from ultrasound scatterers. This chapter draws results from all previous chapters, highlights the conclusions and indicates the limitations of the work in thesis. Suggestions for further research are also discussed.

8.1 Conclusions

Generally speaking, advanced signal processing techniques have not attracted enough attention in ultrasound contrast imaging in ultrasonics. The analysis of measured echo signals from ultrasound scatterers is limited and thus traditional signal processing techniques become a bottleneck in the development of ultrasound contrast imaging. The objective of this thesis has been stated in Chapter 1 and there is a brief recapitulation in the following:

- Extract the characteristics of measured echo signals from ultrasound scatterers in the time and the frequency domains with high accuracy.
- Reveal the true spectra of ultrasound microbubbles based on experimental measurements by removing the influence of ultrasound transducer, which can explore the real behavior of single microbubbles in ultrasound contrast imaging.
- Design a single joint estimation framework to characterise measured ultrasound scatterers automatically, which provides a fundamental basis for discrimination of various ultrasound scatterers.

Consequently, more sophisticated signal processing techniques should be adopted and some existing signal processing algorithms should be improved in the research field of ultrasound contrast imaging. This thesis has introduced and developed several popular and effective advanced signal processing approaches in the analysis of ultrasound echo returns based on actual experimental measurements.

8.1.1 Outcomes of the thesis

- First of all, several popular *non-parametric* signal processing methods employed in other applications, such as audio processing [75, 83] and radar detection [165, 166], have been utilised to characterise ultrasound echo signals in Chapter 3. In the time domain, a new method termed as HTWD-VAD, which combines an envelope detection method and a voice activity detection technique, has been proposed to estimate pulse locations of echo signals; in the frequency domain, multitaper spectrum estimation has been introduced to estimate the spectral content of echo signals. Although these methods have their own limitations, they have been shown to have better performance than simple signal processing methods that are traditionally used in ultrasonics.

- As an alternative to the *non-parametric* methods, Chapter 4 and Chapter 6 perform the analysis using a *parametric* method. Chapter 4 has proposed a signal model, which includes both the temporal and the spectral information for measured ultrasound echo signals. Moreover, modern numerical Bayesian methods, together with reversible jump Markov chain Monte Carlo algorithms, have been developed to obtain the estimates of the model parameters in Chapter 6. The proposed signal model and the developed estimation algorithm can generate more accurate results than those obtained from *non-parametric* methods.

- By combining the *non-parametric* methods and the *parametric* methods, it is possible to achieve a joint estimation system for model parameters in both the time and the frequency domains automatically in the context of a properly proposed signal model. From the signal processing point of view, this proposed approach has suggested a new way to look at the time-frequency analysis using a *parametric* method. From the perspective of ultrasonics, the spectral content of measured ultrasound echo signals can be estimated with high accuracy. The temporal content of the echoes, which are often neglected in the ultrasound literature, can also be determined. Having obtained these spectral and temporal information, it is easier to discriminate ultrasound microbubble echoes from soft tissue when being treated in a single unified framework.

- Considering that all the extracted characteristics of ultrasound echo signals are based on experimental measurements, the influence of the transducer cannot be

ignored. Since the intuitive approach of designing an (inverse) filter to remove the transducer effect cannot produce satisfactory estimation results, the signal model proposed in Chapter 4 has been modified to incorporate transducer characteristics as part of the model parameters, which is presented in Chapter 7. By using Bayesian inference and a modified reversible jump Markov chain Monte Carlo algorithm, which is specially developed for the modified signal model, the true spectral content of ultrasound echo signals has been revealed. This modified signal model and the modified estimation algorithm have provided the real spectra of echoes from ultrasound scatterers. In particular, the real MB behavior without the impact of the transducer has been exploited.

8.1.2 Limitations of the work

The major limitations of the proposed *parametric* estimation approach are threefold from the signal processing point of view.

- With respect to pulse location estimation in the time domain, the number of pulses is determined by a *non-parametric* method. There are two assumptions made in Section 3.2.3. If the length of a pulse segment is smaller than 20 data points, the pulse may not be detected and may be regarded as noise. If two pulse segments in the signal are very close to each other, i.e. the distance between two pulses is smaller than 10 data points, or they overlap with each other, the estimation method cannot distinguish individual pulses and may regard them as one single pulse with a longer width. However, there are some situations in measured ultrasound echo signals that two consecutive pulses or two overlapping pulses are present. For example, Figure 8.1 shows an illustration of two consecutive pulses with different frequency components in a noisy-free signal. Under these insufficient assumptions, the estimated number of pulses in the signal will be incorrect.

- Moreover, the incorrect temporal estimation will also affect the estimation of the spectral content of a signal in the frequency domain. It may detect some spurious frequencies in the pulse or lose some important frequency information. Therefore, a new *parametric* estimation method for estimating the number of pulse segments in a signal is worth being explored in future work.

- Another limitation depends on how closely a proposed signal model matches an

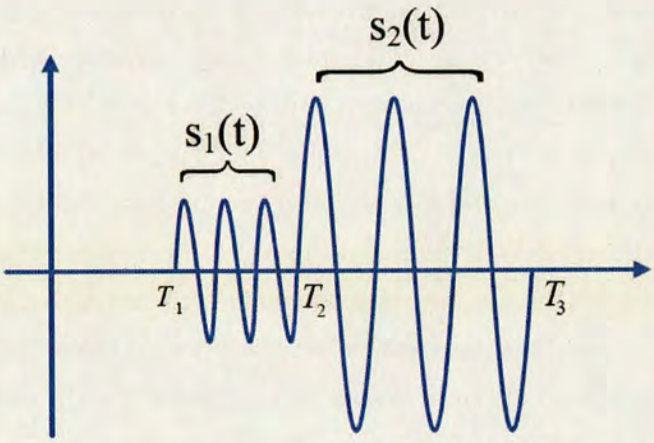


Figure 8.1: *The diagram for a simple illustration of two consecutive pulses.*

experimentally measured echo signal. The mismatch between them can only be reduced rather than being avoided, as long as *parametric* estimation approaches are used. Although the errors between the model and the measurements are inevitable, they can be reduced by having more understandings about the physical procedure that generates the signal and proposing a more appropriate signal model for the measurements.

- Since the estimation approaches used in this thesis are mainly on the basis of statistical signal processing techniques, they require a significant amount of computations. Therefore, it is not easy to implement in some applications, especially when the processing time is a very critical issue.

8.2 Suggestions for future research

The objective of this thesis is to develop statistical signal processing tools for the analysis of measured ultrasound echo signals. Considering the limitations discussed in Section 8.1.2, Section 8.2.1 and Section 8.2.2 aim at proposing some interesting suggestions from the perspective of advancing signal processing techniques. Moreover, Section 8.2.3 and Section 8.2.4 provide some advice for further research from the perspective of improving ultrasound contrast imaging techniques.

8.2.1 Extension work on pulse location estimation

In this thesis, the number of pulse segments in a signal is estimated using a proposed HTWD-VAD method in the time domain, which is *non-parametric*. The limitation of it has already been pointed out in Section 8.1.2.

There is also a *parametric* method that can be used to estimate the number of pulse segments and their corresponding locations. This idea regard a single pulse segment embedded in noise as an object with two change-points. The update of the number of pulse segments can be implemented by incorporating two complementary moves using a rjMCMC algorithm: the *birth* and *death* moves, the *merge* and *split* moves. In this way, the update of pulse locations will convert to the update of each pulse object. Moveover, the two change-points for each pulse object can be updated iteratively in pair. Figure 8.2 shows an illustration of four kinds of moves for updating a single pulse object.

Although this updating process is able to circumvent the limitations stated in Section 8.1.2, it is still computational intensive. This is because there are two sets of parameters that need to be updated in each iteration: one is for updating the number of pulse objects with two change-points, and the other is for updating the number of frequencies in the corresponding pulse object. It is worth being investigated further in spite of high computational complexity.

8.2.2 Extension work on various signal models

In Chapter 4, a lot of discussion about why a sum of sinusoids model for a single pulse segment is chosen has been presented. It only shows an intuitive choice about the signal model, with the consideration of a compromise between the suitability and the complexity. Nevertheless, it does not provide a justification of why this selected model is the most appropriate one among various models. This will result in another broad research area of signal modeling for experimentally measured ultrasound echo signals.

8.2.3 Discrimination of ultrasound scatterers

From the ultrasound contrast imaging point of view, it is of significant importance to differentiate contrast agent echoes from the surrounding soft tissue in order to enhance the quality of ultrasound images. As described in Section 2.2, the experimental measurements use SCSs to mimic tissue behavior and use commercial MBs as UCAs.

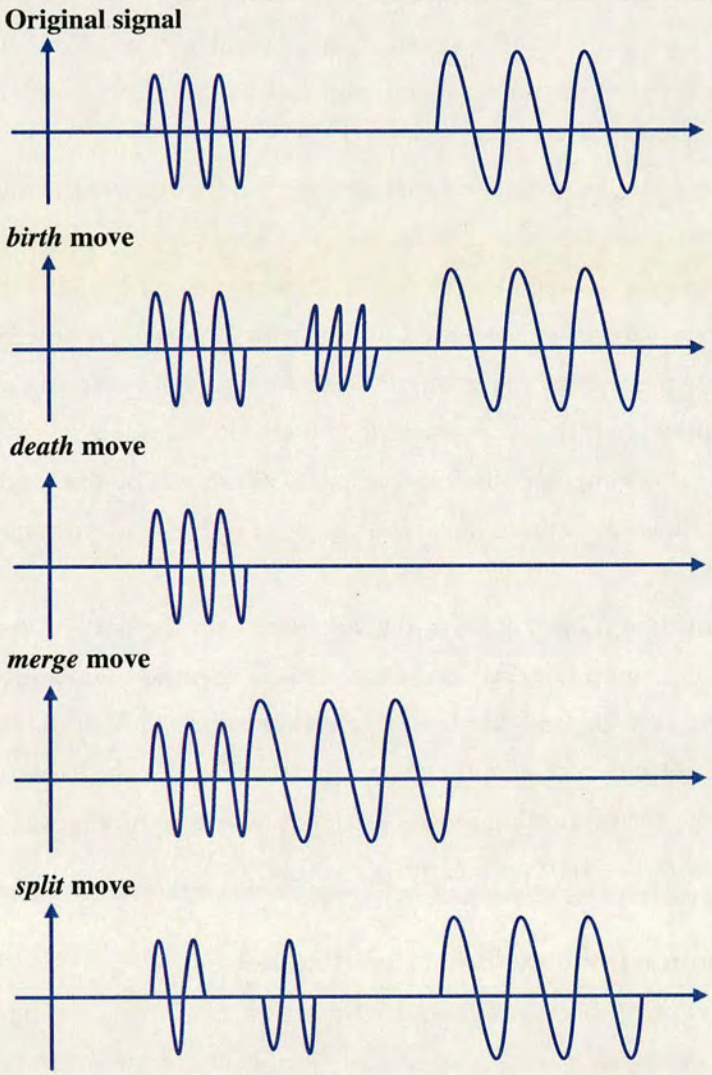


Figure 8.2: *The diagram for a simple illustration of four different moves for updating a pulse object.*

Therefore, discriminating echo signals from SCSs and MBs becomes a more and more encouraging research topic in ultrasound contrast imaging.

In general, discrimination can be categorized into a classification or pattern recognition problem. Pattern recognition is a sub-topic of machine learning, which is a subfield of artificial intelligence [167]. It is concerned with the development of techniques

that allow computers to learn. Research about the pattern recognition can be divided into supervised learning and unsupervised learning. In supervised learning, the objective is to learn the mapping from input data to the outputs which are determined by a supervisor. In unsupervised learning, there is no such a supervisor and only the input data are available. It aims to find the regularities in the input data set. Therefore, some algorithms and learning methods that commonly used in artificial neural networks (ANNs) can be adopted in differentiating ultrasound echo signals.

In the process of solving a classification problem, feature extraction from observed signals is of a significant importance. These features should have obvious difference among distinct signals. Specifically, the feature extraction for measured ultrasound echo signals includes characterising their temporal and spectral contents. The available features can be summarised as follows:

- Pulse length in the time domain
- Bandwidth in the frequency domain
- The number of frequency in a single pulse segment
- Means and variance of frequency values in a single pulse segment
- Total signal energy

All these information can be used as features in classifying measured ultrasound echo signals from SCSs and MBs. However, the pulse length in the time domain and the bandwidth in the frequency domain are very similar for both SCS responses and MB responses. Therefore, they cannot be taken as the input data to the ANN.

8.2.4 Adaptive waveform design

Most of the recent developments in waveform design for the excitation pulses do not use the information of ultrasound echo returns. It is because the conventionally developed excitation schemes only focused on improving CTRs at individual frequency components, which cannot be estimated in a joint probabilistic estimation framework.

Having obtained the extracted characteristics of ultrasound echo returns using the proposed advanced statistical signal processing techniques, it is possible to make the transmit pulse waveform be adapted to the measured echo returns and, make the receiver be adapted to noise statistics and the characteristics of the echo signals. The

adaptive transmit waveform selection and the adaptive receiver techniques are commonly used in radar and communications system design [165], which can also be adopted and extended to the research field of ultrasound contrast imaging in ultrasonics.

References

[1] "Precision acoustics products: the membrane hydrophone." [Online]. Available: http://www.acoustics.co.uk/products

[2] V. Sboros, S. Pye, T. Anderson, C. Moran, M. Averkiou, C. MacDonald, J. Gomatam, and W. McDicken, "The measurement of backscatter from individual contrast agent microbubbles," in *IEEE Ultrasonic Symposium*, vol. 2, 2002, pp. 1945–1947.

[3] V. Sboros, S. D. Pye, C. A. MacDonald, J. Gomatam, C. M. Moran, and W. N. McDicken, "Absolute measurement of ultrasonic backscatter from single microbubbles," *Ultrasound in Med.Biol.*, vol. 31, pp. 1063–1072, 2005.

[4] W. N. McDicken, *Diagnostic ultrasonics: Principles and use of instruments*. John Wiley and Sons, Inc.New York, 1976.

[5] B. B. Goldberg, Ed., *Ultrasound Contrast Agents*. London: Martin Dunitz, 1997.

[6] T. Albrecht and J. Hohmann, "Contrast agents in sonography," *Toshiba VISIONS*, 2004.

[7] H. Barthel, P. Price, and E. O. Aboagye, "Advances in research - small-animal imaging of tumour proliferation with PET," *LANCET ONCOL. 5:100-100.*, 2004.

[8] S. Yang, M. Law, and D. Z. et al., "Dynamic contrast-enhanced perfusion mr imaging measurements of endothelial permeability: Differentiation between atypical and typical meningiomas," *AJNR Am. J. Neuroradiol.*, vol. 24(8), pp. 1554 – 1559, 2003.

[9] T. G. Purdie, E. Henderson, and T. Y. Lee, "Functional CT imaging of angiogenesis in rabbit vx2 soft-tissue tumor," *Phys Med Biol*, vol. 46, pp. 3161–3175, 2001.

[10] E. Stride and N. Saffari, "Microbubble ultrasound contrast agents: a review," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 217, pp. 429–447, 2003.

[11] R. Gramiak and P. M. Shah, "Echocardiography of the aortic root," *Invest Radio*, vol. 3, pp. 356–366, 1968.

[12] A. Bouakaz, J. Borsboom, and N. de Jong, "New contrast imaging method using double frequency exposure," in *IEEE Ultrason.Sympos.Proc.*, vol. 1, 23-27 Aug. 2004, pp. 339–342.

[13] S. Hilgenfeldt, D. Lohse, and M. Zomack, "Sound scattering and localized heat deposition of pulse-driven microbubbles," *Journal of the Acoustical Society of America*, vol. 107(6), pp. 3530–3539, 2000.

[14] N. de Jong, R. Cornet, and C. T. Lancee, "Higher harmonics of vibrating gas filled microspheres, part one:simulations," *Ultrasonics*, vol. 32, pp. 447–453, 1994.

[15] P. J. A. Frinking and N. de Jong, "Acoustic modelling of shell-encapsulated gas bubbles," *Ultrasound Med.Biol.*, vol. 24, pp. 523–533, 1998.

[16] C. C. Church, "The effects of an elastic solid surface layer on the radial pulsations of gas bubbles," *J.Acoustic.Soc.Am.*, vol. 97, pp. 1510–1521, 1995.

[17] K. Morgan, J. Allen, P. Dayton, J. Chomas, A. Klibanov, and K. Ferrara, "Experimental and theoretical ecvaluation of microbubble behavior: Effect of transmitted phase and bubble size," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 47, pp. 1494–1509, 2000.

[18] C. T. Chin, C. Lancee, J. Borsboom, F. Mastik, M. Frijlink, N. de Jong, M. Versluis, and D. Lohse, "A 25 million frames per second digital camera with 128 highly sensitive frames," *Rev.Sci.Instrum.*, vol. 74, pp. 5026–5034, 2003.

[19] J. Borsboom, "Advanced detection strategies for ultrasound contrast agents," Ph.D. dissertation, Interuniversity cardiology institute of the Netherlands, 2005.

[20] R. Hickling and M. S. Plesset, "Collapse and rebound of a spherical bubble in water," *Physics of Fluids*, vol. 7(1), pp. 7–14, 1964.

[21] B. C. Eatock and R. Y. Nishi, "Numerical studies of the spectrum of low intensity ultrasound scattered by bubbles," *Journal of the Acoustical Society of America*, vol. 77, pp. 1692–1701, 1985.

[22] K. Vokurka, "Comparison of rayleigh's, herring's and gilmore's models of gas bubbles," *Acustica*, vol. 59, pp. 214–219, 1986.

[23] N. D. Jong, R. Cornet, and C. T. Lancee, "Higher harmonics of vibrating gas filled microspheres, part one: simulations," *Ultrasonics*, vol. 32, pp. 447–453, 1994.

[24] P. Marmottant, S. van der Meer, M. Emmer, M. Versluis, N. de Jong, S. Hilgenfeldt, and D. Lohse, "A model for large amplitude oscillations of coated bubbles accounting for buckling and rupture," *J.Acoust.Soc.Am.*, vol. 118(6), pp. 3499–3505, 2005.

[25] V. Sboros, K. V. Ramnarine, C. M. Moran, S. D. Pye, and W. N. McDicken, "Understanding the limitations of ultrasonic backscatter measurements from microbubble populations," *Phys Med Biol*, vol. 47, pp. 4287–4299, 2002.

[26] D. O. North, "Analysis of the factors which determines signal/noise discrimination in radar," RCA Laboratories, Princeton, New Jersy, Tech. Rep., 1943.

[27] S. M. Kay, *Fundamentals of statistical signal processing, volume 2: detection theory*. Prentice Hall PTR, 1998.

[28] P. D. Krishna, P. M. Shankar, and V. L. Newhouse, "Subharmonic generation from ultrasonic contrast agents," *Phys Med Biol*, vol. 44, pp. 681–694, 1999.

[29] N. de Jong, L. Hoff, T. Skotland, and N. Bom, "Absorption and scatter of the encapsulated gas filled microspheres: theoretical considerations and some measurements," *Ultrasonics*, vol. 30, pp. 95–103, 1992.

[30] J. N. Marsh, M. S. Hughes, and C. S. H. et al., "Frequency and concentration dependence of the backscatter coefficient of the ultrasound contrast agent," *J. Ac. Soc. Am.*, vol. 104, pp. 1654–1666, 1998.

[31] P. Gregory, *Bayesian inference and maximum entropy methods in science and engineering*. Kluwer academic publishers, 2001.

[32] G. L. Bretthorst, *Bayesian spectrum analysis and parameters estimation*. Springer-Verlag, N.Y., 1988.

[33] C. Andrieu and A. Doucet, "Joint bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Process.*, vol. 47, pp. 2667–2676, 1999.

[34] K. Copsey, N. Gordon, and A. Marrs, "Bayesian analysis of generalized frequency-modulated signals," *IEEE Trans. Signal Process.*, vol. 50, pp. 725–735, 2002.

[35] N. de Jong, P. J. A. Frinking, A. Bouakaz, and F. J. T. Cate, "Detection procedures of ultrasound contrast agents," *Ultrasonics*, vol. 38, pp. 87–92, 2000.

[36] B. A. Schrope and V. L. Newhouse, "Second harmonic ultrasonic blood perfusion measurement," *Ultrasound Med.Biol.*, vol. 19(7), pp. 567–579, 1993.

[37] P. N. Burns, "Instrumentation for contrast echocardiography," *Echocardiography*, vol. 19, pp. 241–258, 2002.

[38] F. Forsberg, W. T. Shi, and B. B. Goldberg, "Subharmonic imaging of contrast agents," *Ultrasonics*, vol. 38, pp. 93–98, 2000.

[39] D. H. Simpson, C. T. Chin, and P. N. Burns, "Pulse inversion doppler: a new method for detecting non-linear echoes from microbubble contrast agents," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 46, pp. 372–382, 1999.

[40] G. Brock-Fisher, M. Poland, P. Rafter, and M. Mooney, "Experimental observations of the sensitivity and frequency response of the power modulation technique for contrast imaging," in *Fifth Heart Center European Symposium on Ultrasound Contrast Imaging*, 2000, pp. 1–77.

[41] R. J. Eckersley, C. T. Chin, and P. N. Burns, "Optimising phase and amplitude modulation schemes for imaging microbubble contrast agents at low acoustic power," *Ultrasound in Med.&Biol.*, vol. 31, pp. 213–219, 2005.

[42] M. Averkiou, M. Bruce, S. Jensen, P. Rafter, T. Brock-Fisher, and J. Powers, "Pulsing shemes for the detection of nonlinear echoes from contrast microbubbles," unpublished.

[43] N. de Jong, "Mechanical index," *Eur. J. Echocard.*, vol. 3:73, 2002.

[44] C. E. Cook and W. M. Siebert, "The early history of pulse compression radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24(4), pp. 825–833, 1988.

[45] T. Misaridis, "Ultrasound imaging using coded signals," Ph.D. dissertation, Technical University of Denmark, 2001.

[46] J. R. Klauder, A. C. Price, S. Darlington, and W. J. Albersheim, "The theory and design of chirp radars," *Bell System Technical Journal*, vol. 39(4), pp. 745–808, 1960.

[47] R. J. Purdy, P. E. Blankenship, C. E. Muehe, C. M. Rader, E. Stern, and R. C. Williamson, "Radar signal processing," *Lincoln laboratory Journal*, vol. 12(2), pp. 297–320, 2000.

[48] R. J. Eckersley, M. X. Tang, K. Chetty, and J. V. Hajnal, "Microbubble contrast agent detection using binary coded pulses," *Ultrasound Med Biol.*, vol. 33(11), pp. 1787–1795, 2007.

[49] J. J. Faran, "Sound scattering by solid cylinders and spheres," *Journal of the Acoustical Society of America*, vol. 23, pp. 405–418, 1951.

[50] R. Hickling, "Analysis of echoes from a solid elastic sphere in water," *Journal of the Acoustical Society of America*, vol. 34, pp. 1582–1593, 1962.

[51] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: Spectral estimation, signal modeling, adaptive filtering and array processing.* McGrawHill Companies, Inc., 2000.

[52] Alencar, S. Marcelo, J. Rocha, and C. Valdemar, *Communication systems*, 2005.

[53] S. L. Hahn, *Hilbert transforms in signal processing.* Artech House Signal Processing Library, 1996.

[54] M. Feldman, "Theoretical analysis and comparison of the hilbert transform decomposition methods," *Mechanical Systems and Signal Processing*, vol. 22, pp. 509–519, 2008.

[55] V. Nagesh, "Application of digital signal processing techniques to partial discharge measurements," Ph.D. dissertation, Indian institute of science, 1993.

[56] S. Sriram, S. Nitin, K. M. M. Prabhu, and M. J. Bastiaans, "Signal denoising techniques for partial discharge measurements," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 12, pp. 1182–1191, 2005.

[57] T. W. Parks and C. S. Burrus, *Digital filter design.* New York: John Wiley and Sons, 1987.

[58] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Process. Mag.*, vol. 9, pp. 21–67, 1992.

[59] D. W. Griffin and J. S. Lim, "Signal estimation from modified short-time fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 236–243, 1984.

[60] M. W. J. van Hest, "Multi resolution analysis of partial discharge signals," Master's thesis, Technische Universiteit Eindhoven, 1999.

[61] N. I. Cho, C. H. Choi, and S. U. Lee, "Adaptive line enhancement by using an IIR lattice notch filter," *IEEE Trans. Acoust. Speech. Signal Process.*, vol. 37, pp. 585–589, 1989.

[62] J. R. Treichler, C. R. J. Jr., and M. G. Larimore, *Theory and design of adaptive filters*. New York: John Wiley and Sons, 1987.

[63] P. M. Clarkson, *Optimal and adaptive signal processing*. CRC Press Inc., 1993.

[64] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, pp. 112–114, 2004.

[65] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proc.Roy.Soc.London*, vol. 460, pp. 1597–1611, 2004.

[66] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[67] Z. Lam, "Wavelets: A tutorial," April 2002, coursenotes.

[68] B. Vidakovic and P. Mueller, "Wavelets for kids: A tutorial introduction," duke University.

[69] A. Graps, "An introduction to wavelets," *IEEE computational science and engineering*, vol. 2(2), pp. 50–61, 1995.

[70] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to wavelets and wavelet transforms*. Prentice-Hall International, 1998.

[71] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrica*, vol. 81(3), pp. 425–455, 1994.

[72] D. L. Donoho, *Progress in wavelet analysis and applications*, 1993, ch. Progress in wavelet analysis and WVD: a ten minute tour, pp. 109–128.

[73] S. Sardy, "Minimax threshold for denoising complex signals with waveshrink," *IEEE Trans. Signal Process.*, vol. 48(4), pp. 1023–1028, 2000.

[74] J. Benesty, M. M. Sondhi, and Y. Huang, Eds., *Speech production*. Springer, 2008, ch. Formant estimation and tracking, pp. 213–226.

[75] J.-H. Chang, N. S. Kim, and S. K. Mitra, "Vioce activity detection based on multiple statistical models," *IEEE Trans. Signal Process.*, 2006.

[76] Y. D. Cho and A. Kondoz, "Analysis and improvement of a statistical model-based voice activity detector," *IEEE Signal processing letters*, vol. 8, pp. 276–278, 2001.

[77] J. Sohn, N. S. Kim, and W. Sung, "A statistical model-based voice activity detection," *IEEE Signal Processing Lett.*, vol. 6(1), pp. 1–3, 1999.

[78] J. Ramirez, J. M. Gorriz, and J. C. Segura, *Robust speech recognition and understanding*, 2007, ch. Voice activity detection:Fundamentals and speech recognition system robustness, pp. 1–22.

[79] D. K. Freeman, G. Cosier, C. B. Southcott, and I. Boyd, "The voice activity detector for the PAN-european digital cellular mobile telephone service," in *ICASSP*, vol. 1, 1989, pp. 369–372.

[80] A. Sangwan and M. C. Chiranth, "VAD techniques for real-time speech transmission on the internet," in *IEEE International conference on High-speed networks and multimedia communications*, 2002, pp. 46–50.

[81] K. Itoh and M. Mizushima, "Environmental noise reduction based on speech/non-speech identification for hearing aids," in *ICASSP*, vol. 1, 1997, pp. 419–422.

[82] Y. D. Cho, K. Al-Naimi, and A. Kondoz, "Imroved statistical voice activity detection based on a smoothed statistical likelihood ratio," in *ICASSP*, 2001, pp. 737–740.

[83] A. Davis, S. Nordholm, and R. Togneri, "Statistical voice activity detection using low-variance spectrum estimation and an adaptive threshold," *IEEE Trans. Speech Audio Process.*, vol. 14, pp. 412–424, 2006.

[84] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of IEEE*, vol. 70, pp. 1055–1096, 1982.

[85] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena," *Terrestrial Magnetism and Atmospheric Electricity*, vol. 3, pp. 13–41, 1898.

[86] R. Blackman and J. Turkey, *The measurement of power spectra*. Dover publications, New York, 1958.

[87] M. S. Bartlett, "Smoothing periodograms for time series with continuous spectra," *Nature*, vol. 161, pp. 686–687, 1948.

[88] P. W. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15(2), pp. 70–76, 1967.

[89] A. Khilko, "Fourier-based spectral analysis with adaptive resolution," 2008. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:0802.1348

[90] S. Haykin, *Adaptive filter theory*. Prentice Hall, 2002.

[91] S. M. Kay, *Modern spectral estimation*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[92] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophysical J.Royal Astro.Soc.*, vol. 33, pp. 347–366, 1973.

[93] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem appraoch," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31(10), pp. 1235–1248, 1983.

[94] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT-a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34(4), pp. 1340–1342, 1986.

[95] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37(7), pp. 984–995, 1989.

[96] E. T. Jaynes, *Maximum entropy and Bayesian spectral analysis and estimation problems*, 1987, ch. Bayesian spectrum and Chirp analysis, pp. 1–37.

[97] J. Jensen, S. H. Jensen, and E. Hansen, "Exponential sinusoidal modeling of transitional speech segments," in *Acoustics, Speech, and Signal Processing, 1999. ICASSP '99. Proceedings., 1999 IEEE International Conference on*, vol. 1, Mar 1999, pp. 473–476.

[98] M. Abe and J. O. Smith, "Am/fm rate estimation for time-varying sinusoidal modeling," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, vol. 3, Mar 2005, pp. 201–204.

[99] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian data analysis*. Chapman&Hall/CRC, 2004.

[100] D. Gamerman, *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*. Chapman & Hall, 1997.

[101] D. V. Lindley, "The bayesian approach," *Scandinavian Journal of Statistics*, vol. 5, pp. 1–26, 1978.

[102] A. F. M. Smith, "Bayesian statistics: Present position and potential developments, some personal views with discussion," *Journal of the Royal Statistical Society*, vol. 147, pp. 245–259, 1984.

[103] B. Efron, "Why isn't everyone a bayesian?" *American Statistician*, vol. 40, pp. 1–11, 1986.

[104] R. E. Kass and D. Steffey, "Approximate bayesian inference in conditionally independent hierarchical models (parametric empirical bayes models)," *Journal of the American Statistical Association*, vol. 84, pp. 717–726, 1989.

[105] J. H. Albert, "Computational methods using a bayesian hierarchical generalized linear model," *Journal of the American Statistical Association*, vol. 83, pp. 1037–1045, 1988.

[106] ——, "A MCMC algorithm to fit a general exchangeable model," *Communications in Statistics-Simulation and Computation*, vol. 25, pp. 573–592, 1996.

[107] A. Jeffreys, *The theory of Probability*. Cambridge University Press, 1961.

[108] J. J. K. O. Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian methods applied to signal processing*. Springer, 1996.

[109] J. J. Rajan, "Time series classification," Ph.D. dissertation, Engineering department, Cambridge University, 1994.

[110] P. Stoica and Y. Selen, "Model oeder selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, pp. 36–47, 2004.

[111] S. M. Kay, *Fundamentals of statistical signal processing, volume 1: estimation theory*. Prentice Hall PTR, 1993.

[112] H. Akaike, "Fitting autoregreesive model for prediction," *Annals Institute of statistical mathematics*, vol. 21, pp. 243–247, 1969.

[113] S. M. Kay and S. L. Marple, "Spectrum analysis : A morden perspective," in *Proceedings of IEEE*, vol. 69, no. 11, November 1981, pp. 1380–1420.

[114] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 716–723, 1974.

[115] R. J. Bhansali and D. Y. Downham, "Some properties of the order of an autoregressive model selected by a generalization of akaike's fpe criterion," *Biometrika*, vol. 64, pp. 547–551, 1977.

[116] M. B. Priestley, *Spectral analysis and time series*. Academic Press, London, 1981.

[117] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist*, vol. 6, pp. 461–464, 1978.

[118] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.

[119] E. J. Hannan and B. G. Quinn, "The determination of the order of an autoregression," *Journal of the royal statistical society series B.*, vol. 41, pp. 190–195, 1979.

[120] E. Parzen, "Some recent advances in time series modelling," *IEEE Trans. Autom. Control*, vol. 19, pp. 723–730, 1974.

[121] T. E. Landers and R. T. Lacoss, "Some geophysical applications of autoregressive spectral estimates," *IEEE Trans. Geosci. Electron.*, vol. GE-15, pp. 661–675, 1977.

[122] P. J. Green, "Reversible jump MCMC computation and bayesian model determination," *Biometrika*, vol. 82, pp. 711–732, 1995.

[123] C. C. Heyde and I. M. Johnstone, "On asymptotic posterior normality of stochastic processes," *Journal of the Royal Statistical Society*, vol. 41, pp. 184–189, 1979.

[124] R. E. Kass, L. Tierney, and J. B. Kadane, *Bayesian Statistics 3*. Oxford University Press, 1988, ch. Asymptotic in Bayesian computation (with discussion), pp. 261–278.

[125] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes.* Oxford : OUP, 2001.

[126] C. P. Robert and G. Casella, *Monte Carlo statistical methods.* New York: Springer-Verlag, 1999.

[127] C. Andrieu, "On sequential simulation-based methods for bayesian filtering," University of British Columbia, Tech. Rep., 1998.

[128] C. Andrieu, N. D. Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine learning*, vol. 50, pp. 5–43, 2003.

[129] W. Ng, "Advancres in wideband array signal processing using numerical Bayesian methods," Ph.D. dissertation, McMaster University, 2003.

[130] J. Geweke, "Bayesian inference in econometric models using monte carlo integration," *Econometrica*, vol. 24, pp. 1317–1399, 1989.

[131] R. Y. Rubinstein, Ed., *Simulation and the Monte Carlo method.* John Wiley and Sons (New York), 1981.

[132] C. G. Bucher, "Adaptive sampling-an iterative fast monte carlo procedure," *Structural safety*, vol. 5, pp. 119–126, 1988.

[133] D. Remondo, R. Srinivasnan, and et al., "Adaptive importance sampling for performance evaluation and parameter optimization of communications systems," *IEEE Trans. Comput.*, vol. 48(4), pp. 557–565, 2000.

[134] D. B. Rubin, *Bayesian statistics 3.* Oxford University Press, 1998, ch. Using the SIR algorithm to simulate posterior distributions, pp. 395–402.

[135] N. R. Metropolis, A. W. Rosenbluth, and E. T. et al., "Equation of state calculations by fast computing machine," *Journal of Chemical Physics*, vol. 21, pp. 1087–91, 1953.

[136] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[137] L. Tierney, "Markov chains for exploring posterior distributions," *Amer.Statist.*, vol. 22, pp. 1702–1762, 1994.

[138] C. Andrieu, L. A. Breyer, and A. Doucet, "Convergence of simulated annealing using foster-lyapunov criteria," Cambrige University Engineering Department, Tech. Rep., 1999.

[139] C. Andrieu, N. de Freitas, and A. Doucet, *Robust full Bayesian methods for neural networks.* MIT press, 2000, ch. Advances in neural information processing systems 12, pp. 379–385.

[140] S. Geman and D. Geman, "Stochastic relaxation,gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, pp. 721–732, 1984.

[141] A. Doucet and X. Wang, "Monte carlo methods for signal processing: A review in the statistical signal processing context," *IEEE Signal Process. Mag.*, pp. 152–170, 2005.

[142] L. Dou and R. J. W. Hodgson, "Bayesian inference and gibbs sampling in spectral analysis and parameter estimation i," *Inverse Problems*, vol. 11, pp. 1069–1085, 1995.

[143] ——, "Bayesian inference and gibbs sampling in spectral analysis and parameter estimation ii," *Inverse Problems*, vol. 12, pp. 121–137, 1996.

[144] S. P. Meyn and R. L. Tweedie, "Computable bounds for convergence rates of markov chains," *Annals of Applied Probability*, vol. 4, pp. 124–148, 1994.

[145] G. O. Roberts and N. G. Polson, "A note on the geometric convergence of the gibbs sampler," *Journal of the Royal Statistical Society, Series B*, vol. 56, pp. 377–384, 1994.

[146] G. O. Roberts and R. L. Tweedie, "Geometric convergence and central limit theorems, for multidimensional hastings and metropolis algorithms," University of Cambridge, Tech. Rep., 1994.

[147] M. K. Cowles and B. P. Carlin, "Markov chain monte carlo convergence diagnostics: a comparative review," *Journal of the American Statistical Association*, vol. 91, pp. 883–904, 1996.

[148] C. P. Robert, "Convergence control methods for markov chain monte carlo algorithms," *Statistical Science*, vol. 10, pp. 231–253, 1995.

[149] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association*, vol. 85, pp. 398–409, 1990.

[150] S. P. Brooks and G. O. Roberts, "Diagnosing convergence of markov chain monte carlo algorithms," Statistical Laboratory, University of Cambridge, Tech. Rep., 1995.

[151] J. Liu, W. H. Wong, and A. Kong, "Correlation structure and convergence rate of the gibbs sampler: applications to the comparison of estimators and augmentation schemes," *Biometrika*, vol. 81, pp. 27–40, 1994.

[152] G. O. Roberts and S. K. Sahu, "Updating schemes, correlation structure, blocking and parametrization for the gibbs sampler," *Journal of the Royal Statistical Society, Series B*, vol. 59, pp. 291–317, 1997.

[153] R. E. Kass and E. H. Slate, *Bayesian Statistics 4*. Oxford University Press, 1992, ch. Reparametrization and diagnostics of posterior nonnormality (with discussion), pp. 289–305.

[154] C. Jennison, "Discussion of the meeting on gibbs sampling and other markov chain monte carlo methods," *Journal of the Royal Statistical Society, Series B*, vol. 55, pp. 54–56, 1993.

[155] J. Besag and P. J. Green, "Spatial statistics and bayesian computation (with discussion)," *Journal of Royal Statistical Society, Series B*, vol. 55, pp. 25–37, 1993.

[156] U. Grenander and M. Miller, "Representations of knowledge in complex systems (with discussion)," *J. R. Statist. Soc. B*, vol. 56, pp. 549–603, 1994.

[157] B. P. Carlin and S. Chib, "Bayesian model choice via markov chain monte carlo," *J.R.Statist.Soc.B*, vol. 57, pp. 473–484, 1995.

[158] S. Godsill, "On the relationship between MCMC model uncertainty methods," Cambridge University, Engineering Department, England, Tech. Rep., 1998.

[159] ——, "On the relationship between MCMC model uncertainty methods," *Journal of Computational and Graphical Statistics*, vol. 10(2), pp. 230–248, 2001.

[160] S. L. Lauritzen, *Graphical models*. Oxford university press, 1996.

[161] S. Richardson and P. J. Green, "On bayesian analysis of mixtures with an unknown number of components," *J.R.Statist.Soc.,ser.B*, vol. 59, pp. 731–792, 1997.

[162] C. Andrieu, "MCMC methods for bayesian analysis of nonlinear parametric regression problems: applications to spectral analysis and impulsive deconvolution," Ph.D. dissertation, Univ.Cergy-Pontoise, Cergy-Pontoise, France, 1999.

[163] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, *Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers, 2000, ch. Rao blackwellised particle filtering for dynamic Bayesian networks, pp. 176–183.

[164] W. R. Gilks and G. O. Roberts, *Markov chain Monte Carlo in practice*. Chapman & Hall, 1996, ch. Strategies for improving MCMC, pp. 89–110.

[165] S. P. Sira, D. Cochran, and A. P.-S. et. al., "Adaptive waveform design for improved detection of low-rcs targets in heavy sea clutter," *IEEE Journal of selected topics in signal processing*, vol. 1, pp. 56–66, 2007.

[166] S. Suvurova and S. B. Howard, "Waveform libraries for radar tracking applications: maneuvering targets," in *40th Annual Conf.Information Sciences and Systems*, 2006, pp. 1424–1428.

[167] E. Alpaydin, *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*. MIT press, 2004.

[168] R. W. Daniels, *Approximation methods for electronic filter design (with applications to passive, active, and digital networks)*. McGraw-Hill, 1974.

Appendix A
Generate a simulated signal

The signal has a total length of 1500 data points. There are two pulse segments in this simulated signal. The first pulse locates between 450 and 600, and the second pulse locates between 750 and 850, which are denoted in data points. Each single pulse has different amplitudes and frequency components. A single pulse segment $i = 1, 2$ in the synthetic signal can be constructed as follows:

$$s_i(t) = \sum_{j=1}^{k_i} a_{c_{j,k_i}} \cos(\omega_{j,k_i} t) + a_{s_{j,k_i}} \sin(\omega_{j,k_i} t), \tag{A.1}$$

where k_i represents the number of frequency components in the ith pulse of the signal. $a_{c_{j,k_i}}$ and $a_{s_{j,k_i}}$ denote the amplitudes for a single frequency ω_{j,k_i}. With regard to each sinusoid ω_{j,k_i}, energy $E_{j,k_i} \triangleq a^2_{c_{j,k_i}} + a^2_{s_{j,k_i}}$ is defined. Moreover, the phase information for ω_{j,k_i} can be represented by $\phi_{j,k_i} \triangleq -\arg\tan(a_{s_{j,k_i}}/a_{c_{j,k_i}})$. Since there are two pulse segments in the signal, the parameters setup for each pulse are shown in Table A.1 and Table A.2.

j	E_{j,k_1}	$-\arg\tan(a_{s_{j,k_1}}/a_{c_{j,k_1}})$	ω_{j,k_1}
1	20	0	0.6π
2	20	$\pi/4$	0.7π

Table A.1: *Parameters setup for the 1^{st} pulse segment in the constructed synthetic signal*

j	E_{j,k_2}	$-\arg\tan(a_{s_{j,k_2}}/a_{c_{j,k_2}})$	ω_{j,k_2}
1	40	$\pi/3$	0.2π
2	40	$\pi/5$	0.3π
3	40	$\pi/6$	0.32π

Table A.2: *Parameters setup for the 2^{nd} pulse segment in the constructed synthetic signal*

The simulated corrupting noise is zero-mean identically independent distributed Gaussian noise with a variance of $\sigma^2 = 0.76$. According to the SNR definition for multiple pulse signal in Chapter 3, the $\text{SNR}_1 = 7.1\text{dB}$ and $\text{SNR}_2 = 12.9\text{dB}$ for the two pulses respectively. Finally, the averaged $\overline{\text{SNR}} = \text{SNR}_1 + \text{SNR}_2 = 20\text{dB}$ is obtained.

Appendix B
Derivation of the posterior distribution of the parameters in Chapter 6

B.1 Deriving the posterior distribution of the parameters in a single pulse segment

This section describes the simplification of the posterior distribution of the parameters for a single pulse segment i, including the integration of amplitudes $\mathbf{a}_{k_i}$ and noise variance $\sigma_{k_i}^2$. Then the derivations and the integrations are extended to the full length signal model in Section B.2.

According to (6.4a) and (4.20), the posterior distribution can be obtained directly using Bayes's rule:

$$\begin{aligned} p(k_i, \mathbf{a}_{k_i}, \boldsymbol{\omega}_{k_i}, \sigma_{k_i}^2, \Lambda_i, \delta_i^2 | \mathbf{y}_i) \quad \propto \quad & \left(\frac{1}{2\pi\sigma_{k_i}^2}\right)^{L_i/2} \cdot \exp\left[-\frac{(\mathbf{y}_i - \boldsymbol{D}_i\mathbf{a}_{k_i})^T(\mathbf{y}_i - \boldsymbol{D}_i\mathbf{a}_{k_i})}{2\sigma_{k_i}^2}\right] \\ \times \quad & \frac{1}{|2\pi\sigma_{k_i}^2 \boldsymbol{\Sigma}_{k_i}|^{1/2}} \cdot \exp\left[-\frac{\mathbf{a}_{k_i}^T \boldsymbol{\Sigma}_{k_i}^{-1} \mathbf{a}_{k_i}}{2\sigma_{k_i}^2}\right] \cdot \frac{1}{\pi^{k_i}} \cdot \frac{\Lambda_i^{k_i}/k_i!}{e^{\Lambda_i}} \\ \times \quad & \frac{1}{\sigma_{k_i}^{2(v_i/2+1)}} \cdot \exp\left[-\frac{\gamma_i}{2\sigma_{k_i}^2}\right] \cdot \Lambda_i^{\epsilon_1 - 1/2} \exp[-\epsilon_2 \Lambda_i] \\ \times \quad & \frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right]. \end{aligned} \tag{B.1}$$

The power of the exponentials including $\sigma_{k_i}^2$ can be rearranged together as below:

$$\begin{aligned} \mathbf{Q} &\triangleq (\mathbf{y}_i - \boldsymbol{D}_i\mathbf{a}_{k_i})^T(\mathbf{y}_i - \boldsymbol{D}_i\mathbf{a}_{k_i}) + \mathbf{a}_{k_i}^T\boldsymbol{\Sigma}_{k_i}^{-1}\mathbf{a}_{k_i} + \gamma_i \\ &= (\mathbf{y}_i^T - \mathbf{a}_{k_i}^T\boldsymbol{D}_i^T)(\mathbf{y}_i - \boldsymbol{D}_i\mathbf{a}_{k_i}) + \mathbf{a}_{k_i}^T\boldsymbol{\Sigma}_{k_i}^{-1}\mathbf{a}_{k_i} + \gamma_i \\ &= \mathbf{y}_i^T\mathbf{y}_i - \mathbf{a}_{k_i}^T\boldsymbol{D}_i^T\mathbf{y}_i - \mathbf{y}_i^T\boldsymbol{D}_i\mathbf{a}_{k_i} + \mathbf{a}_{k_i}^T(\boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1})\mathbf{a}_{k_i} + \gamma_i \\ &= \mathbf{y}_i^T\mathbf{y}_i - \mathbf{a}_{k_i}^T\boldsymbol{D}_i^T\mathbf{y}_i - \mathbf{y}_i^T\boldsymbol{D}_i\mathbf{a}_{k_i} + \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} + \gamma_i. \end{aligned}$$

Define $\boldsymbol{M}_{k_i}^{-1} \triangleq \boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1} = (1+\delta_i^{-2})\boldsymbol{D}_i^T\boldsymbol{D}_i$, then $(\boldsymbol{M}_{k_i}^{-1})^T = [(1+\delta_i^{-2})\boldsymbol{D}_i^T\boldsymbol{D}_i]^T = \boldsymbol{M}_{k_i}^{-1}$. Accordingly, $\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1} = \boldsymbol{I}_{2k_i}$.

Therefore, $\mathbf{Q}$ can be represented as:

$$
\begin{aligned}
\mathbf{Q} &= \mathbf{y}_i^T\mathbf{y}_i + \mathbf{y}_i^T\boldsymbol{D}_i(\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1} - \boldsymbol{I}_{2k_i})\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i + \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} \\
&\quad - \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} - \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i + \gamma_i \\
&= \mathbf{y}_i^T\mathbf{y}_i - \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y} + \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} - \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} - \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i \\
&\quad + \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i + \gamma_i \\
&= \mathbf{y}_i^T\mathbf{y}_i - \mathbf{y}_i^T\boldsymbol{D}_i(\boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1})^{-1}\boldsymbol{D}_i^T\mathbf{y}_i + \mathbf{a}_{k_i}^T(\boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1})\mathbf{a}_{k_i} \\
&\quad - \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\mathbf{a}_{k_i} - \mathbf{a}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i \\
&\quad + \mathbf{y}_i^T\boldsymbol{D}_i\boldsymbol{M}_{k_i}^T\boldsymbol{M}_{k_i}^{-1}\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i + \gamma_i.
\end{aligned}
$$

Define:

$$
\begin{aligned}
\boldsymbol{P}_{k_i} &\triangleq \boldsymbol{I}_{L_i} - \boldsymbol{D}_i\boldsymbol{M}_{k_i}\boldsymbol{D}_i^T && \text{(B.2)} \\
\boldsymbol{M}_{k_i}^{-1} &\triangleq \boldsymbol{D}_i^T\boldsymbol{D}_i + \boldsymbol{\Sigma}_{k_i}^{-1} && \text{(B.3)} \\
\boldsymbol{m}_{k_i} &\triangleq \boldsymbol{M}_{k_i}\boldsymbol{D}_i^T\mathbf{y}_i, && \text{(B.4)}
\end{aligned}
$$

then

$$
\mathbf{Q} = \mathbf{y}_i^T\boldsymbol{P}_{k_i}\mathbf{y}_i + (\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i})^T\boldsymbol{M}_{k_i}^{-1}(\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i}) + \gamma_i. \qquad \text{(B.5)}
$$

Therefore, (B.1) can be rearranged using definitions in (B.2), (B.3) and (B.4) and the

posterior distribution becomes:

$$
\begin{aligned}
p(k_i, \mathbf{a}_{k_i}, \boldsymbol{\omega}_{k_i}, \sigma_{k_i}^2, \Lambda_i, \delta_i^2 | \mathbf{y}_i) \quad \propto \quad & (\frac{1}{2\pi\sigma_{k_i}^2 \boldsymbol{\Sigma}_{k_i}})^{1/2} \cdot \exp\left[-\frac{(\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i})^T \boldsymbol{M}_{k_i}^{-1} (\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i})}{2\sigma_{k_i}^2} \right] \\
\times \quad & \left(\frac{1}{2\pi\sigma_{k_i}^2} \right)^{L_i/2} \cdot \exp\left[-\frac{\mathbf{y}_i^T \boldsymbol{P}_{k_i} \mathbf{y}_i}{2\sigma_{k_i}^2} \right] \\
\times \quad & (\sigma_{k_i}^2)^{-v_i/2-1} \cdot \exp\left[-\frac{\gamma_i}{2\sigma_{k_i}^2} \right] \\
\times \quad & \frac{(\Lambda_i/\pi)^{k_i}}{k_i!} \cdot \frac{\Lambda_i^{\epsilon_1 - 1/2} e^{-\epsilon_2 \Lambda_i}}{e^{\Lambda_i}} \\
\times \quad & \frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2} \right] .
\end{aligned} \tag{B.6}
$$

After the rearrangement of the posterior distribution $p(k_i, \mathbf{a}_{k_i}, \boldsymbol{\omega}_{k_i}, \sigma_{k_i}^2, \Lambda_i, \delta_i^2 | \mathbf{y}_i)$, the integrations of amplitude coefficients $\mathbf{a}_{k_i}$ and the noise variance $\sigma_{k_i}^2$ become much easier.

Since the prior for amplitude coefficients $\mathbf{a}_{k_i}$ is a multivariate Gaussian distribution and it is chosen as a conjugate prior, the posterior distribution is also a multivariate Gaussian distribution with mean vector $\boldsymbol{m}_{k_i}$ and covariance matrix $\sigma_{k_i}^2 \boldsymbol{M}_{k_i}$. The expression of it is given as:

$$
\begin{aligned}
p(\mathbf{a}_{k_i} | k_i, \boldsymbol{\omega}_{k_i}, \sigma_{k_i}^2, \Lambda_i, \delta_i^2) = & \frac{1}{(2\pi)^{2k_i/2} \cdot |\sigma_{k_i}^2 \boldsymbol{M}_{k_i}|^{1/2}} \\
& \times \exp\left[-\frac{(\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i})^T \boldsymbol{M}_{k_i}^{-1} (\mathbf{a}_{k_i} - \boldsymbol{m}_{k_i})}{2\sigma_{k_i}^2} \right] ,
\end{aligned} \tag{B.7}
$$

The joint posterior distribution after integration out $\mathbf{a}_{k_i}$ is presented:

$$
\begin{aligned}
p(k_i, \boldsymbol{\omega}_{k_i}, \sigma_{k_i}^2, \Lambda_i, \delta_i^2 | \mathbf{y}_i) \quad \propto \quad & \left(\frac{1}{2\pi\sigma_{k_i}^2} \right)^{L_i/2} \cdot \exp\left[-\frac{1}{2\sigma_{k_i}^2} (\gamma_i + \mathbf{y}_i^T \boldsymbol{P}_{k_i} \mathbf{y}_i) \right] \\
\times \quad & \left(\frac{1}{1+\delta_i^2} \right)^{k_i} \cdot (\sigma_{k_i}^2)^{-v_i/2-1} \cdot \frac{(\Lambda_i/\pi)^{k_i}}{k_i!} \cdot \frac{\Lambda_i^{\epsilon_1 - 1/2} e^{-\epsilon_2 \Lambda_i}}{e^{\Lambda_i}} \\
\times \quad & \frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2} \right] .
\end{aligned} \tag{B.8}
$$

Similarly, the prior for the nuisance parameter $\sigma_{k_i}^2$ is also a conjugate prior. The prior

and the posterior are both inverse Gamma distributions. According to an investigation of (B.6), the posterior distribution of $\sigma^2_{k_i}$ can be denoted as $\sigma^2_{k_i}|k_i, \boldsymbol{\omega}_{k_i}, \Lambda_i, \delta_i^2, \mathbf{y}_i \sim \mathcal{IG}\left(\frac{L_i+v_i}{2}, \frac{\gamma_i+\mathbf{y}_i\boldsymbol{P}_{k_i}\mathbf{y}_i}{2}\right)$, and the expression of it is given as:

$$\begin{aligned} p(\sigma^2_{k_i}|k_i, \boldsymbol{\omega}_{k_i}, \Lambda_i, \delta_i^2, \mathbf{y}_i) &= \frac{\left(\gamma_i + \mathbf{y}_i^T\boldsymbol{P}_{k_i}\mathbf{y}_i\right)^{(L_i+v_i)/2}}{\Gamma(\frac{L_i+v_i}{2})} \cdot (\sigma^2_{k_i})^{-\frac{L_i+v_i}{2}-1} \\ &\times \exp\left(-\frac{\gamma_i + \mathbf{y}_i^T\boldsymbol{P}_{k_i}\mathbf{y}_i}{2\sigma^2_{k_i}}\right), \end{aligned} \tag{B.9}$$

where $\Gamma(\cdot)$ is a Gamma function. As a result, the simplified posterior distribution after integration of $\sigma^2_{k_i}$ can be achieved as follows:

$$\begin{aligned} p(k_i, \boldsymbol{\omega}_{k_i}, \Lambda_i, \delta_i^2|\mathbf{y}_i) &\propto \left(\frac{1}{1+\delta_i^2}\right)^{k_i} \cdot \left(\frac{1}{\gamma_i + \mathbf{y}_i^T\boldsymbol{P}_{k_i}\mathbf{y}_i}\right)^{(L_i+v_i)/2} \cdot \frac{1}{k_i!} \cdot \left(\frac{\Lambda_i}{\pi}\right)^{k_i} \\ &\times \frac{\Lambda_i^{\epsilon_1-1/2}e^{-\epsilon_2\Lambda_i}}{e^{\Lambda_i}} \cdot \frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right]. \end{aligned} \tag{B.10}$$

The posterior distribution shown in (B.10) is the same one as shown in (6.6).

B.2 Deriving the posterior distribution of the parameters in a full length signal model

In order to update the pulse locations in the time domain, i.e. the change-point positions, the posterior distribution of all parameters for a full length signal model can be achieved using (4.18) and (6.31) following Bayes's rule.

Similar to (B.1) in Section B.1, the posterior distribution based on the full length

signal model can be extended to include all m pulse segments, which is expressed as:

$$
\begin{aligned}
&p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \mathbf{a}_{k_m}, \boldsymbol{\omega}_{k_m}, \sigma_n^2, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2 | \mathbf{y}) \\
&\qquad \propto \left(\frac{1}{2\pi\sigma_n^2}\right)^{N/2} \cdot \exp\left[-\frac{(\mathbf{y}-\mathbf{D}\mathbf{a}_{k_m})^T(\mathbf{y}-\mathbf{D}\mathbf{a}_{k_m})}{2\sigma_n^2}\right] \\
&\qquad \times \frac{1}{|2\pi\sigma_n^2\boldsymbol{\Sigma}_{k_m}|^{1/2}} \cdot \exp\left[-\frac{\mathbf{a}_{k_m}^T\boldsymbol{\Sigma}_{k_m}^{-1}\mathbf{a}_{k_m}}{2\sigma_n^2}\right] \\
&\qquad \times \frac{1}{\sigma_n^{2(v/2+1)}} \cdot \exp\left[-\frac{\gamma}{2\sigma^2}\right] \\
&\qquad \times \prod_{i=1}^{m}\left\{\frac{1}{\pi^{k_i}} \cdot \frac{\Lambda_i^{k_i}/k_i!}{e^{\Lambda_i}} \times \Lambda_i^{\epsilon_1-1/2} \cdot \exp[-\epsilon_2\Lambda_i]\right\} \\
&\qquad \times \prod_{i=1}^{m}\left\{\frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right]\right\} \\
&\qquad \times \left(\frac{1}{N-1} \cdot \frac{1}{N-2} \cdots \frac{1}{N-2m}\right).
\end{aligned} \tag{B.11}
$$

Moreover, in analogous to the derivation in Section B.1, the posterior distribution in (B.11) can be rearranged in (B.12) so as to make the integrations of $\mathbf{a}_{k_m}$ and σ_n^2 a lot easier in subsequent analysis.

$$
\begin{aligned}
&p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \mathbf{a}_{k_m}, \boldsymbol{\omega}_{k_m}, \sigma_n^2, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2 | \mathbf{y}) \\
&\qquad \propto \left(\frac{1}{2\pi\sigma_n^2\boldsymbol{\Sigma}_{k_m}}\right)^{1/2} \cdot \exp\left[-\frac{(\mathbf{a}_{k_m}-\boldsymbol{m}_{k_m})^T\boldsymbol{M}_{k_m}^{-1}(\mathbf{a}_{k_m}-\boldsymbol{m}_{k_m})}{2\sigma_n^2}\right] \\
&\qquad \times \left(\frac{1}{2\pi\sigma_n^2}\right)^{N/2} \cdot \exp\left[-\frac{\mathbf{y}^T\boldsymbol{P}_{k_m}\mathbf{y}}{2\sigma_n^2}\right] \\
&\qquad \times (\sigma_n^2)^{-v/2-1} \cdot \exp\left[-\frac{\gamma}{2\sigma_n^2}\right] \\
&\qquad \times \prod_{i=1}^{m}\left\{\frac{(\Lambda_i/\pi)^{k_i}}{k_i!} \cdot \frac{\Lambda_i^{\epsilon_1-1/2}e^{-\epsilon_2\Lambda_i}}{e^{\Lambda_i}}\right\} \\
&\qquad \times \prod_{i=1}^{m}\left\{\frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right]\right\}.
\end{aligned} \tag{B.12}
$$

This posterior distribution is also very complicated and the nuisance parameters $\mathbf{a}_{k_m}$ and σ_n^2 can be integrated out, which are similar to the integrations in (B.7) and (B.9). The first step is the integration of $\mathbf{a}_{k_m}$ and the posterior distribution after this

integration is given as:

$$
\begin{aligned}
&p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \sigma_n^2, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2 | \mathbf{y}) \\
&\qquad \propto \left(\frac{1}{2\pi\sigma_n^2}\right)^{(N/2)} \cdot \exp\left[-\frac{1}{2\sigma_n^2}(\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m}\mathbf{y})\right] \cdot (\sigma_n^2)^{-v/2-1} \\
&\qquad \times \prod_{i=1}^{m} \left\{ \left(\frac{1}{1+\delta_i^2}\right)^{k_i} \cdot \frac{(\Lambda_i/\pi)^{k_i}}{k_i!} \cdot \frac{\Lambda_i^{\epsilon_1 - 1/2} e^{-\epsilon_2 \Lambda_i}}{e^{\Lambda_i}} \right\} \\
&\qquad \times \prod_{i=1}^{m} \left\{ \frac{1}{\delta_i^{2(\alpha_{\delta_i^2}+1)}} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right] \right\}.
\end{aligned} \tag{B.13}
$$

The second step is the integration of σ_n^2, which has an inverse Gaussian distribution. It can be denoted as:

$$
\sigma_n^2 | \boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2, \mathbf{y} \sim \mathcal{IG}\left(\frac{N+v}{2}, \frac{\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m}\mathbf{y}}{2}\right). \tag{B.14}
$$

Moreover, the expression of the posterior distribution of σ_n^2 is presented:

$$
\begin{aligned}
p(\sigma_n^2 | \boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2, \mathbf{y}) &= \frac{\left(\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m}\mathbf{y}\right)^{(N+v)/2}}{\Gamma(\frac{N+v}{2})} \cdot (\sigma_n^2)^{-\frac{N+v}{2}} \\
&\times \exp\left(-\frac{\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m}\mathbf{y}}{2\sigma_n^2}\right),
\end{aligned} \tag{B.15}
$$

where $\Gamma(\cdot)$ is a Gamma function.

Accordingly, the posterior distribution after the integration of σ_n^2 can be obtained:

$$
\begin{aligned}
p(\mathbf{T}_{2m}, \boldsymbol{k}_m, \boldsymbol{\omega}_{k_m}, \boldsymbol{\Lambda}_{1:m}, \boldsymbol{\delta}_{1:m}^2 | \mathbf{y}) &\propto (\gamma + \mathbf{y}^T \boldsymbol{P}_{k_m}\mathbf{y})^{-(N+v)/2} \\
&\times \prod_{i=1}^{m} \left\{ \left(\frac{1}{1+\delta_i^2}\right)^{2k_i} \delta_i^{-2(\alpha_{\delta_i^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta_i^2}}{\delta_i^2}\right] \right\} \\
&\times \prod_{i=1}^{m} \left\{ \frac{1}{k_i!} \cdot \left(\frac{\Lambda_i}{\pi}\right)^{k_i} \times \Lambda_i^{\epsilon_1 - 1/2} \cdot e^{-\Lambda_i(1+\epsilon_2)} \right\}
\end{aligned} \tag{B.16}
$$

This posterior distribution shown in (B.16) is the same one as shown in (6.34).

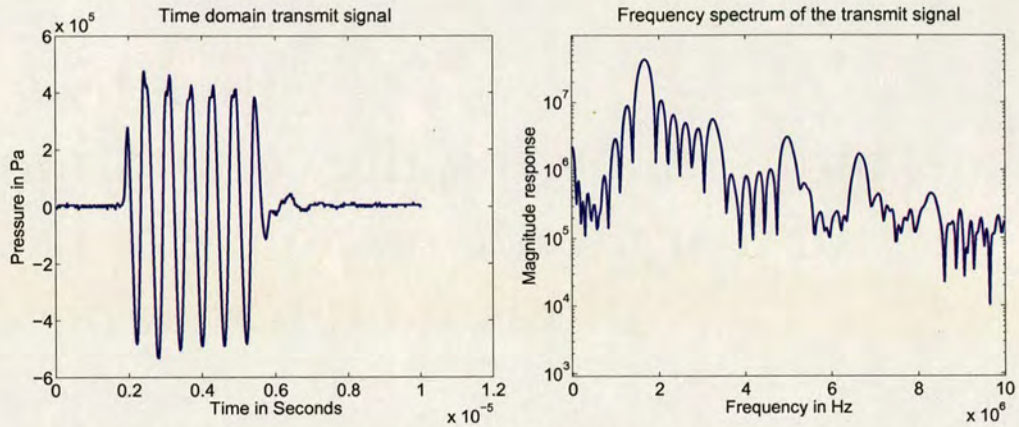
Appendix C
An example showing the calibration of magnitude response of the ultrasound transducer

Take a specific transmit frequency at $f_{xmit} = 1.67$MHz and a SCS with a radius of 49μm as an example to give details about the ultrasound receiver calibration. In this example, fundamental frequency of the transmit pulse is $f_0 = 1.67$MHz, and the second harmonic is $2f_0 = 3.34$MHz. In general, there are four steps in calculating a single calibration point in the spectrum of the ultrasound receiver with the transmit frequency of $f_{xmit} = 1.67$MHz. They will be discussed in the following sections.

C.1 Filtering measured transmit signal

Figure C.1(a) shows the transmit pulse in the time domain. Its frequency spectrum is displayed in Figure C.1(b). It can be seen that most of the energy of the signal concentrates on the fundamental frequency, $f_0 = 1.67$MHz. If the fundamental component and the second harmonic component of the signal are required to be extracted, a band-pass filter is necessary.

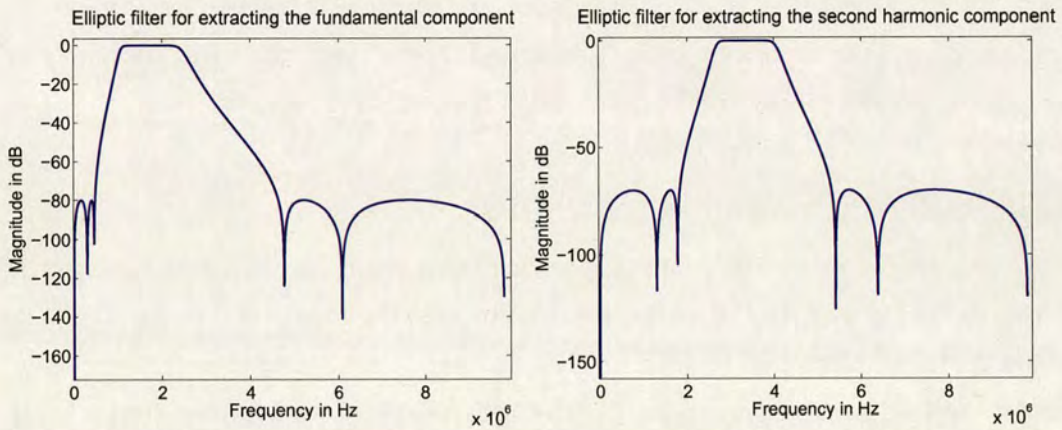
There are two issues needs to be considered. One is determining which type of filter can be used and the other is how to choose a proper pass band of the filter. First of all, among all the popular linear filters, the elliptic filters offer steeper roll-off characteristics than all others, such as Butterworth or Chebyshev filters, if the same number of coefficients are used. However, they show ripples on the whole bandwidth. Moreover, in general, elliptic filters meet given performance specifications with the lowest order of any filter type [168]. In addition to choosing the filter types, the pass band, $f_{passband}$, of the filter should be constrained to: $f_{passband} < \triangle f$, in which $\triangle f$ is the difference between the second harmonic frequency and the fundamental frequency. The fundamental frequencies of the transmit signal varies from 1.2MHz to 4.0MHz, and thereby the minimum difference between the second harmonic and the fundamental frequency amongst them is $\triangle f_{min} = 1.26$MHz. As a consequence, the pass



(a) Transmit pulse in the time domain.

(b) Frequency spectrum of the transmit pulse.

Figure C.1: *Transmit pulse with the fundamental frequency at* 1.67*MHz.*



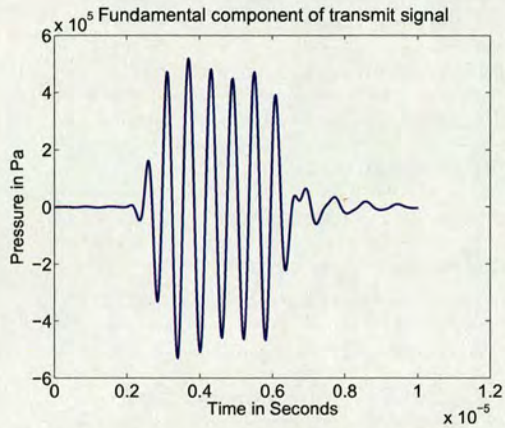
(a) Frequency spectrum of the elliptic filter for extracting the fundamental frequency component.

(b) Frequency spectrum of the elliptic filter for extracting the second harmonic frequency component.

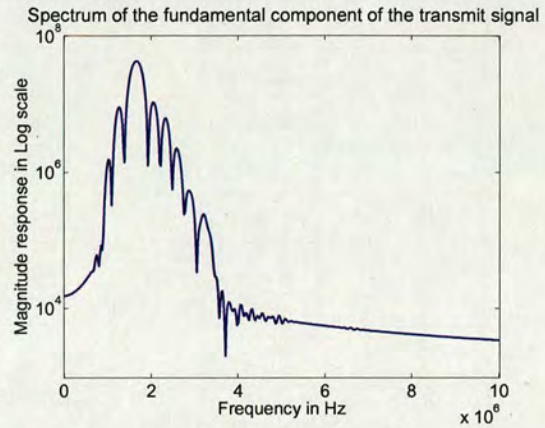
Figure C.2: *Frequency spectra of the elliptic filters for extracting the fundamental and the second harmonic frequency components.*

band is set to $f_{passband} = 0.8$MHz heuristically, as suggested in [3]. If the selected $f_{passband} > \triangle f_{min}$, the extracted frequency components will be overlapped, which may result in spectrum distortion. Figure C.2(a) and Figure C.2(b) show the frequency spectra of the elliptic filters that are used to extract the fundamental frequency component and the second harmonic component of the transmit signal, and the measured SCS echo signal. The filters have their pass bands centering at $f_0 = 1.67$MHz and $2f_0 = 3.34$MHz respectively.

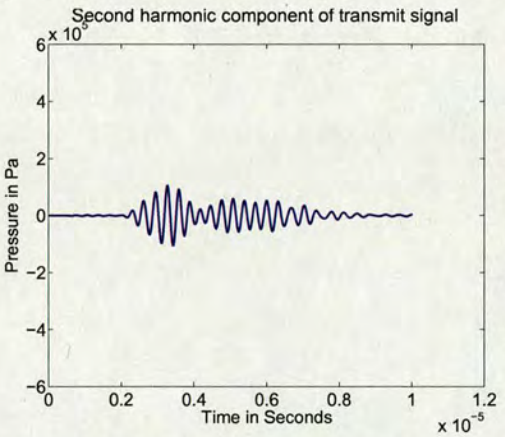
Using two elliptic filters illustrated in Figure C.2, the original measured transmit



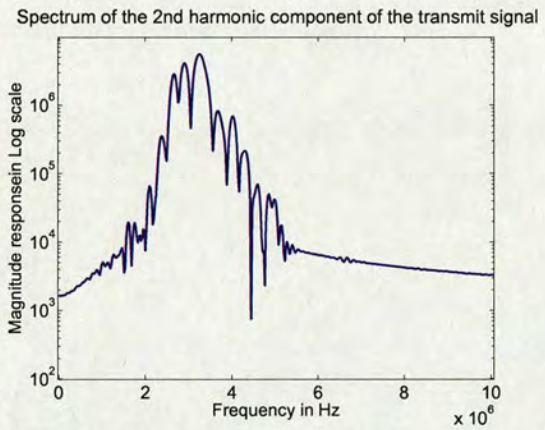
(a) The fundamental component of the transmit pulse in the time domain.



(b) The frequency spectrum of the fundamental component of the transmit pulse.



(c) The second harmonic component of the transmit pulse in the time domain.



(d) The frequency spectrum of the second harmonic component of the transmit pulse.

Figure C.3: *The fundamental and the second harmonic components of the transmit pulse in the time and the frequency domains.*

pulse can be filtered to produce its fundamental component and the second harmonic component in Figure C.3(a) and Figure C.3(c). Seen from the time domain signals, it is shown that the energy of transmit pulse almost focuses on its fundamental component. Moreover, their frequency spectra are displayed in Figure C.3(b) and Figure C.3(d) respectively.

C.2 Generating the theoretical SCS responses

There are many theoretical descriptions about the SCS behavior available in ultrasonics literature [13,50]. One of the most widely used analytical expression of the pressure of

Parameters	Values	Physical meanings
c	1480m/s	velocity of sound in water
c_1	3921.67m/s	velocity of compressional waves in copper
c_2	2321.77m/s	velocity of shear waves in copper
R_0	49μm	radius of copper sphere
r_0	0.075m	distance between the center of the copper sphere and the surface of the transducer
k	ω/c	wave number in water
k_1	ω/c_1	compressional wave number in water
k_2	ω/c_2	shear wave number in water

Table C.1: *Parameters and their physical meanings in (C.1).*

theoretical backscattered pulse from SCSs in the time domain is stated in [3,50] in the form of: [1]

$$P_{scatt}(t) = \frac{R_0}{2r_0\sqrt{2\pi}} \times \int_{\omega_{min}}^{\omega_{max}} \left\{ g(\omega)\alpha(\omega) f_\infty(k, k_1, k_2) e^{j\omega\frac{r_0}{c}} \right\} e^{j\omega t} d\omega, \tag{C.1}$$

where

$$g(\omega) = \int_{-\infty}^{\infty} P_I(t) e^{-j\omega t} dt. \tag{C.2}$$

Moreover, the physical meanings and values of all the parameters in (C.1) are listed in Table C.1.

Furthermore, there is an illustration of physical interpretation of (C.1), which is shown in Figure C.4. In general, $P_I(t)$ is the measured transmit pulse in the time domain, and $P_{scatt}(t)$ is the calculated theoretical backscattered pulse in the time domain. All operations are carried out in the frequency domain and then the SCS scattering, $P_{scatt}(t)$, is obtained by transforming back to the time domain using the inverse FT. $g(\omega)$ denotes the FT of the transmit pulse $P_I(t)$ in the frequency domain. $\alpha(\omega)$ is a water attenuation function. It depends on the spectral component:

$$\alpha(\omega) = \exp\{-\alpha(r_0 - R_0)(\frac{\omega}{2\pi})^2\}, \tag{C.3}$$

in which α is an attenuation coefficient at 20^oC. The SCS scattering is given by the

[1] j in this equation represents the imaginary unit in the complex number system.

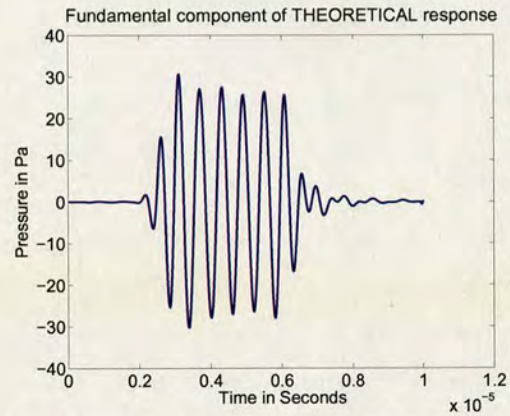
Figure C.4: *Physical interpretation of the procedure of obtaining the scattered wave pulse from an SCS.*

function, $f_\infty(k, k1, k2)$, which can only be evaluated using a lookup table. In this function, $k = \frac{\omega}{c}$, $k_1 = \frac{\omega}{c_1}$, $k_2 = \frac{\omega}{c_2}$. The values in the table are related to a **Ka** value, which is defined as: $\mathbf{Ka} = \frac{\omega}{c} \times R_0$, in [50]. It ranges from 0 to $12R_0$ according to measurements in [3]. In order to ensure all frequency components in the transmit pulse will be taken into account in the theoretically calculated backscattered pulse, the frequency limits ω_{min} and ω_{max} need to be set within the **Ka** value range. Therefore, in the experimental measurements in [3], the sampling frequency of the transmit pulses is $f_s = 50\text{MHz}$, i.e., $f_{max} = 50\text{MHz}$, $\mathbf{Ka}_{max} = \frac{2\pi}{c} f_{max} R_0 = 11.8871 R_0 < 12R_0$, such that all the frequencies in the transmit pulse can be considered. Moreover, $e^{j\omega\frac{r_0}{c}}$ is related to a retard function, which only gives a certain time delay.

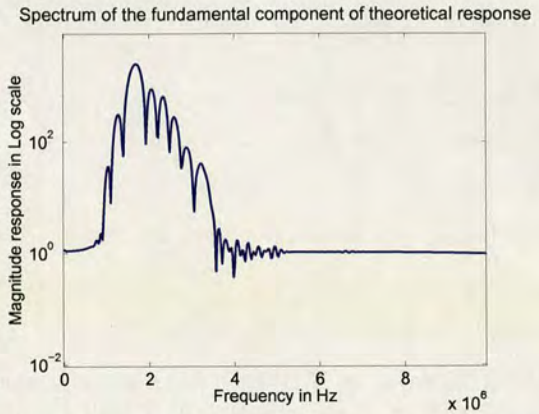
Using the theoretical calculation, both the fundamental frequency and the second harmonic of the transmit pulse can produce their corresponding backscattering pulse in the time domain respectively, which are displayed in Figure C.5(a) and Figure C.5(c). Moreover, their frequency spectra are also shown in Figure C.5(b) and Figure C.5(d). Compared to the wave shapes and the pulse durations to the transmit pulse in Figure C.3(a) and Figure C.3(c), they are almost the same, yet with attenuated amplitudes.

C.3 Filtering the measured SCS response

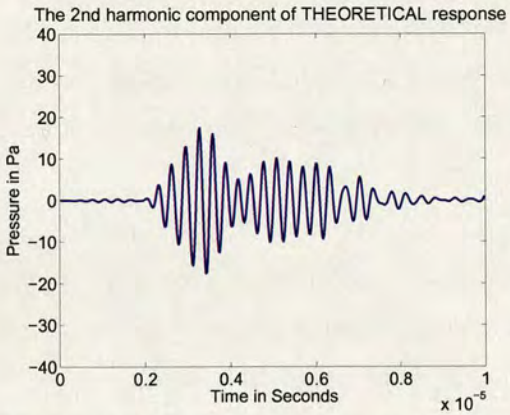
Apart from the measured transmit pulse, filtering the measured ultrasound SCS response is also necessary. They are filtered using the same elliptic filters presented in Section C.1. Figure C.6(a) and Figure C.6(c) show the filtered ultrasound SCS echo signals in the time domain. The pulse durations of the filtered fundamental component and the filtered second harmonic component are almost the same as the generated theoretical SCS response, however, the shapes of the waves do not look similar. Their respective frequency spectra shown in Figure C.6(b) and Figure C.6(d) are not of the similar shape either. The reason is that the scattered waves in measurements are subjected to the spectral characteristics and phase alterations of the ultrasound receiver [3]. That is also the way how the receiver is calibrated.



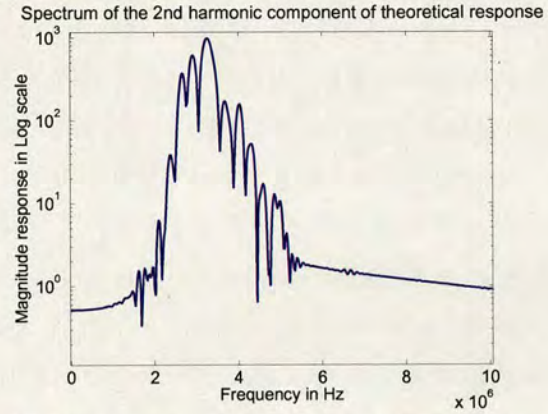
(a) The calculated fundamental component of the theoretical SCS response in the time domain.



(b) The frequency spectrum of the fundamental component of the theoretical SCS response.



(c) The second harmonic component of the transmit pulse in the time and the frequency domains.



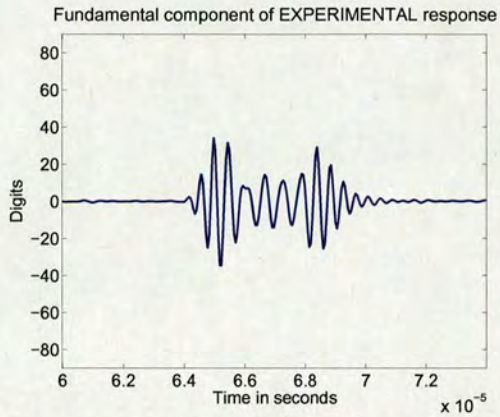
(d) The frequency spectrum of the second harmonic component of the theoretical SCS response.

Figure C.5: *The fundamental and the second harmonic components of the theoretical SCS response in the time and the frequency domains.*

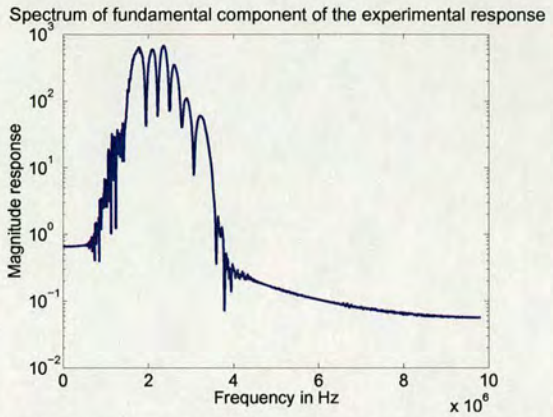
C.4 Calculating the ratio of energy density of the measured response to the theoretical response

Having obtained the fundamental and the second harmonic components of both the theoretical and experimentally measured SCS responses, a calibration point in the frequency spectrum can be calculated as:

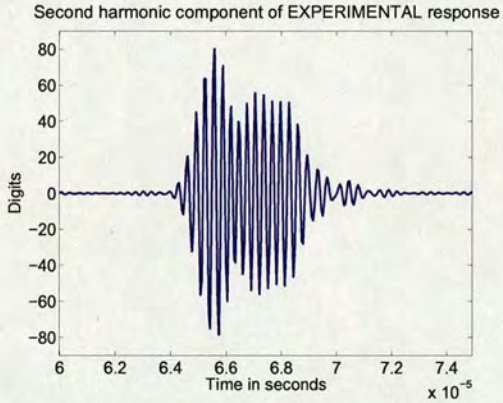
$$G^2_{calibrate_{mag}}(f_{mag}) = \frac{\text{Energy Density}_{exp}}{\text{Energy Density}_{theory}} = \frac{\int_{t^{(exp)}_{min}}^{t^{(exp)}_{max}} |P_{exp}(t)|^2 dt}{\int_{t^{(theory)}_{min}}^{t^{(theory)}_{min}} |P_{theory}(t)|^2 dt}, \tag{C.4}$$



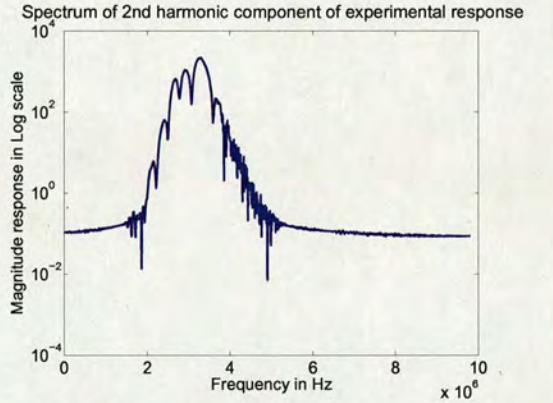
(a) The fundamental component of the measured SCS response in the time domain.



(b) The frequency spectrum of the fundamental component of the measured SCS response.



(c) The second harmonic component of the measured SCS response in the time domain.



(d) The frequency spectrum of the second harmonic component of the measured SCS response.

Figure C.6: *The fundamental and the second harmonic component of the measured SCS response in the time and the frequency domains.*

where $P_{exp}(t)$ and $P_{theory}(t)$ represent the backscattered radiation waveform in the time domain. The time intervals, i.e. $[t_{min}^{(theory)}, t_{max}^{(theory)}]$ and $[t_{min}^{(exp)}, t_{max}^{(exp)}]$, over which the integration are evaluated should contain sufficient duration to include the complete pulse segment for both the theoretical responses and the measured responses. As a result, at the transmit frequency of 1.67MHz, two calibration points in the frequency spectrum of the calibrated ultrasound receiver at $f_0 = 1.67$MHz and $2f_0 = 3.34$MHz can be obtained using (C.4).

Appendix D

Derivation of posterior distribution of model parameters in Chapter 7

If $0 < k \leqslant k_{max}$, the complicated joint posterior distribution is expressed as:

$$\begin{aligned} p(k, \boldsymbol{\psi}_k | \mathbf{y}_i, 0 < k \leqslant k_{max}) \propto & (2\pi\sigma_k^2)^{-L/2} \cdot \exp\left\{-\frac{1}{2\sigma_k^2}[\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k]^T[\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k]\right\} \\ & \times \frac{\Lambda^{k+\epsilon_1-1/2}}{k!}\exp[-\Lambda(1+\epsilon_2)] \cdot \frac{1}{|2\pi\sigma_k^2\boldsymbol{\Sigma}_k|^{1/2}} \cdot \exp\left[-\frac{\boldsymbol{a}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{a}_k}{2\sigma_k^2}\right] \\ & \times \frac{1}{\pi^k} \cdot \frac{\exp[-\gamma/(2\sigma_k^2)]}{\sigma_n^{2(v/2+1)}} \cdot (\delta^2)^{-\alpha_{\delta^2}-1} \cdot \exp\left(-\frac{\beta_{\delta^2}}{\delta^2}\right) \\ & \times p(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_p). \end{aligned} \tag{D.1}$$

Rearrange the power of the exponential with respect to σ_k^2 as follows:

$$\begin{aligned} \mathbf{Z} &\triangleq (\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k)^T(\mathbf{y}_i - \mathbf{H}\boldsymbol{D}\boldsymbol{a}_k) + \boldsymbol{a}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{a}_k + \gamma \\ &= \mathbf{y}_i^T\mathbf{y} - \boldsymbol{a}_k^T\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i - \mathbf{y}_i^T\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k + (1+\delta^{-2})\boldsymbol{a}_k^T\boldsymbol{D}^T\mathbf{H}^T\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k + \gamma. \end{aligned} \tag{D.2}$$

Define $\boldsymbol{R}^{-1} \triangleq (1+\delta^{-2})\boldsymbol{D}^T\mathbf{H}^T\mathbf{H}\boldsymbol{D}$, and $(\boldsymbol{R}^{-1})^T = [(1+\delta^{-2})\boldsymbol{D}^T\mathbf{H}^T\mathbf{H}\boldsymbol{D}]^T = \boldsymbol{R}^{-1}$, which results in $\boldsymbol{R}^T\boldsymbol{R}^{-1} = \boldsymbol{I}_{2k}$. Therefore, $\mathbf{Z}$ can be expressed as:

$$\begin{aligned} \mathbf{Z} &= \mathbf{y}_i^T\mathbf{y}_i - \boldsymbol{a}_k^T\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i - \mathbf{y}_i^T\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k + \boldsymbol{a}_k^T\boldsymbol{R}^{-1}\boldsymbol{a}_k + \gamma \\ &= \mathbf{y}_i^T\mathbf{y}_i + \mathbf{y}_i^T\mathbf{H}\boldsymbol{D}(\boldsymbol{R}^T\boldsymbol{R}^{-1} - \boldsymbol{I}_{2k})\boldsymbol{R}\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i + \boldsymbol{a}_k^T\boldsymbol{R}^{-1}\boldsymbol{a}_k \\ &\quad -\mathbf{y}_i^T\mathbf{H}\boldsymbol{D}(\boldsymbol{R}^T\boldsymbol{R}^{-1})\boldsymbol{a}_k - \boldsymbol{a}_k^T(\boldsymbol{R}^{-1}\boldsymbol{R})\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i + \gamma \\ &= \mathbf{y}^T\mathbf{y}_i - \mathbf{y}_i^T\mathbf{H}\boldsymbol{D}\boldsymbol{R}\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i + \boldsymbol{a}_k^T\boldsymbol{R}^{-1}\boldsymbol{a}_k \\ &\quad -\mathbf{y}_i^T\mathbf{H}\boldsymbol{D}\boldsymbol{a}_k - \boldsymbol{a}_k^T\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i \\ &\quad +\mathbf{y}_i^T\mathbf{H}\boldsymbol{D}\boldsymbol{R}(\mathbf{H}\boldsymbol{D})^T\mathbf{y}_i + \gamma. \end{aligned} \tag{D.3}$$

Define:

$$\boldsymbol{P}_k^{(r)} \triangleq \boldsymbol{I}_{L\times L} - \mathbf{H}\boldsymbol{D}\boldsymbol{R}\boldsymbol{D}^T\mathbf{H}^T \tag{D.4a}$$

$$\left\{\boldsymbol{M}_k^{(r)}\right\}^{-1} = \boldsymbol{R}^{-1} \triangleq (1+\delta^{-2})\boldsymbol{D}^T\mathbf{H}^T\mathbf{H}\boldsymbol{D} \tag{D.4b}$$

$$\boldsymbol{m}_k^{(r)} = \boldsymbol{r}_k \triangleq \boldsymbol{R}\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i = \left\{\boldsymbol{M}_k^{(r)}\right\}^{-1}\boldsymbol{D}^T\mathbf{H}^T\mathbf{y}_i, \tag{D.4c}$$

and $\mathbf{Z}$ can be simplified as:

$$\mathbf{Z} = \mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i + (\boldsymbol{a}_k - \boldsymbol{r}_k)^T\boldsymbol{R}^{-1}(\boldsymbol{a}_k - \boldsymbol{r}_k) + \gamma. \tag{D.5}$$

As a result, the joint posterior distribution shown in (D.1) can be rearranged in the form of:

$$\begin{aligned} p(\boldsymbol{\psi}_k|\mathbf{y}_i, 0<k\leqslant k_{max}) \propto\ & \frac{1}{(2\pi)^{L/2}|\sigma_k^2\boldsymbol{R}|^{1/2}}\cdot\exp\left[-\frac{1}{2\sigma_k^2}(\boldsymbol{a}_k-\boldsymbol{r}_k)^T\boldsymbol{R}^{-1}(\boldsymbol{a}_k-\boldsymbol{r}_k)\right] \\ & \times\left[\left(\frac{1}{1+\delta^2}\right)^{2k}\right]^{1/2}\cdot\exp\left(-\frac{\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i}{2\sigma_k^2}\right)\cdot\exp\left(-\frac{\gamma}{2\sigma_k^2}\right)\cdot\frac{1}{\sigma_k^{1+L}} \\ & \times\frac{\beta_{\delta^2}}{\Gamma(\alpha_{\delta^2})}\cdot\left(\frac{1}{\delta^2}\right)^{\alpha_{\delta^2}+1} \\ & \times\frac{\Lambda^k/k!}{\pi^k}\cdot\Lambda^{\epsilon_1-1/2}\cdot\exp(-\Lambda-\epsilon_2\Lambda)\cdot\sigma_k^{-2(v/2+1)}\cdot\exp\left(-\frac{\gamma}{2\sigma_k^2}\right) \\ & \times p(\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p). \end{aligned} \tag{D.6}$$

D.1 Integrate out amplitudes $\boldsymbol{a}_k$

The joint posterior distribution has been rearranged in (D.6), which will make it easier to be simplified. Seen from the rearranged equation, the posterior distribution for the vector of amplitudes is a multivariate Gaussian distribution with mean $\boldsymbol{r}_k$ and variance $\sigma_k^2\boldsymbol{R}$:

$$\boldsymbol{a}_k|\mathbf{y}_i \sim \mathcal{N}(\boldsymbol{r}_k, \sigma_k^2\boldsymbol{R}) \tag{D.7}$$

$$\begin{aligned} p(\boldsymbol{a}_k|k,\boldsymbol{\omega}_k,\Lambda,\delta^2,\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p,\mathbf{y}_i) =\ & \frac{1}{(2\pi)^{2k/2}|\sigma_k^2\boldsymbol{R}|^{1/2}} \\ & \times\ \exp\left[-\frac{1}{2\sigma_k^2}(\boldsymbol{a}_k-\boldsymbol{r}_k)^T\boldsymbol{R}^{-1}(\boldsymbol{a}_k-\boldsymbol{r}_k)\right]. \end{aligned} \tag{D.8}$$

By analytically integrating out $\boldsymbol{a}_k$ when $0 < k \leqslant k_{max}$, the resulting posterior

distribution can be achieved as:

$$\begin{aligned} p(k,\boldsymbol{\omega}_k,\sigma_k^2,\Lambda,\delta^2,\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p|\mathbf{y}_i) \propto & \frac{1}{\sigma_k^{1+L}}\cdot\exp\left(-\frac{\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i}{2\sigma_k^2}\right)\cdot\frac{\beta_{\delta^2}}{\Gamma(\alpha_{\delta^2})}\cdot\left(\frac{1}{\delta^2}\right)^{\alpha_{\delta^2}+1} \\ & \times\exp\left(-\frac{\beta_{\delta^2}}{\delta^2}\right)\cdot\frac{\Lambda^k/k!}{\pi^k}\cdot\Lambda^{\epsilon_1-1/2}\cdot\exp[-\Lambda(1+\epsilon_2)] \\ & \times\sigma_k^{-2(v/2+1)}\cdot\exp\left(-\frac{\gamma}{2\sigma_k^2}\right)\cdot\left[\left(\frac{1}{1+\delta^2}\right)^{2k}\right]^{1/2} \\ & \times p(\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p). \end{aligned} \tag{D.9}$$

D.2 Integrate out noise term σ_k^2

Moreover, the posterior distribution for noise term σ_k^2 has an inverse Gamma distribution, when $0 \leqslant k \leqslant k_{max}$. It can be expressed as:

$$\sigma_k^2|\mathbf{y}_i \sim \mathcal{IG}\left(\frac{1+L+v}{2},\frac{\gamma+\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i}{2}\right) \tag{D.10}$$

$$\begin{aligned} p(\sigma_k^2|k,\boldsymbol{\omega}_k,\Lambda,\delta^2,\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p,\mathbf{y}_i) = & \frac{(\frac{\gamma+\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i}{2})^{(1+L+v)/2}}{\Gamma\left(\frac{1+L+v}{2}\right)}\cdot\left(\frac{1}{\sigma_k^2}\right)^{(\frac{1+L+v}{2})+1} \\ & \times\exp\left[-\frac{\gamma+\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i}{2\sigma_k^2}\right]. \end{aligned} \tag{D.11}$$

In this way, the complexity of the joint posterior distribution can be further reduced by integration of the noise term σ_k^2, which gives:

$$\begin{aligned} p(k,\boldsymbol{\omega}_k,\Lambda,\delta^2,\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p|\mathbf{y}_i) \propto & (\gamma+\mathbf{y}_i^T\boldsymbol{P}_k^{(r)}\mathbf{y}_i)^{-(1+L+v)/2}\times\frac{(\Lambda/((\delta^2+1)\pi))^k}{k!} \\ & \times\Lambda^{\epsilon_1-1/2}\cdot\exp[-\Lambda(1+\epsilon_2)]\cdot\delta^{-2(\alpha_{\delta^2}+1)}\cdot\exp\left[-\frac{\beta_{\delta^2}}{\delta^2}\right] \\ & \times p(\boldsymbol{\varphi}_m,\boldsymbol{\varphi}_p). \end{aligned} \tag{D.12}$$

This simplified joint posterior distribution is in the same form as the one displayed in (7.17).

Appendix E
Publications

E.1 Journal Papers

- Yan Yan, James R. Hopgood and Vassilis Sboros, "A Novel Estimation System for Multiple-pulse Echo Signals From Ultrasound Scatterers," *to be submitted to IEEE Transactions on Biomedical Engineering* (on page 232)

- Yan Yan, James R. Hopgood and Vassilis Sboros, "Statistical spectral analysis of echo signals from ultrasound contrast microbubbles," *to be submitted to Ultrasound in Medicine and Biology* (on page 242)

E.2 Conference Papers

- Yan Yan, James R. Hopgood and Vassilis Sboros, "Statistical spectrum analysis for echo signals from microbubbles and solid spheres," *In proceedings of IEEE International Ultrasonic Symposium (IUS), pp.1382-1385, Beijing, China, Nov., 2008* (on page 264)

- Yan Yan, James R. Hopgood and Vassilis Sboros, "A novel estimation system for multiple pulse echo signals from ultrasound contrast microbubbles," *In proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.601-604, Las Vegas, Nevada, USA, April, 2008* (on page 268)

- Yan Yan, James R. Hopgood and Vassilis Sboros, "Analysis of echo signal from single ultrasound contrast microbubble using a reversible jump MCMC algorithm," *In proceedings of IEEE 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Socienty (EMBC), pp.1273-1276, Lyon, France, Aug. 2007* (on page 272)

The original publications are included in the following pages.

A novel estimation system for echo signals from measured ultrasound scatterers

Yan Yan, *Student Member, IEEE,* James R. Hopgood, *Member, IEEE,* and Vassilis Sboros

***Abstract*—The understanding and exploitation of non-linear microbubble signals is an active research area that aims to advance contrast ultrasound into a high sensitivity and specificity diagnostic imaging modality. To discriminate the difference between echoes from tissue and contrast microbubbles, it is of significance to extract as much information of the reflected signals as possible, especially their locations in the time domain and their corresponding spectral content in the frequency domain. In this paper, a novel estimation system for extracting the information of interest is proposed. This estimation technique is based on *non-parametric* methods for coarse estimation, followed by a *parametric* method within a Bayesian framework for estimation refinement. The results show that the pulse location and frequency content can be accurately estimated simultaneously and automatically. This may assist in the design of transmit pulsing regimes in future work.**

***Index Terms*—ultrasound contrast agents, microbubble, Bayesian inference, Markov chain Monte Carlo, parametric model**

I. Introduction

ULTRASOUND Contrast Agents (UCAs) were first discovered accidentally by cardiologist Charles Joiner in the 1960's. However, microbubbles (MBs) have been widely used as UCAs in bio-medical research area since the 1990's [1]. They are composed of gas-filled encapsulated microspheres, usually with diameter below $7\mu m$, that can go through microcirculations in the human body. Compared to soft tissue, the microbubbles are more compressible and expandable when exposed to an oscillating acoustic signal as they have a non-linear acoustic signature [2]. In order to design a transmit pulse that can maximise the difference between responses from microbubbles and tissue and increase the contrast-to-tissue ratio (CTR) [3], it is of particular interest to jointly extract the characteristics of responses from UCAs and tissue in both the time and the frequency domains.

A. Limitations in spectral estimation

Most traditional frequency estimation techniques in ultrasonics are based on the Fourier transform (FT) [2]. However, these carry some inherent limitations [4]. First, it is known that traditional *non-parametric* methods suffer from the trade-off between the duration of sidelobes and the width of spectral peak. The sidelobes are associated to the duration of the signal (rectangular window function), which is usually very short for echo signals. Due to the limited frequency resolution, the FT can not detect some frequencies that may have important physical meanings, or may provide false spectral contents.

Second, As the peaks in the spectrum are not narrow enough, it is not easy to determine the exact positions of peaks if the FT is only used. Various optimisation methods, such as thresholding, needed to be considered to determine the accuracy to which the peak frequency can be determined. By contrast, the Bayesian posterior probability for this signal frequency provides the information directly [5]. Furthermore, it is not possible to determine the unknown number of frequency components.

Third, the FT does not localize in time whereas in ultrasonics analysis, the pulse locations and durations are also of significant importance. Bayesian inference can be easily extended to incorporate the time domain information into parameter estimation in an estimation system.

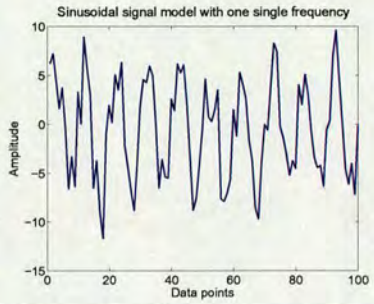
Therefore, a *parametric* model within a Bayesian framework is introduced in this paper in order to improve the quality of frequency estimation. A simulated sinusoidal signal is taken as an example to demonstrate the superiority of Bayesian inference. The signal has a single frequency at 0.2π, which is displayed in Fig. 1(a). It is demonstrated in Fig. 1(b) that the spurious noise features are suppressed and the width of the peak is much narrower than the peak in the DFT-based frequency spectrum. Furthermore, the variances of two peaks can be estimated numerically. They are calculated based on the definition of -3dB bandwidth. According to this definition, the estimated variance for the peak in frequency spectrum, which is shown in the upper panel of Fig. 1(b), is computed as: $Var_{DFT}(f) = 0.0442$; the estimated variance for the peak in the bottom panel in Fig. 1(b) is computed as: $Var_{BI}(f) = 6.0281 \times 10^{(-4)}$. In this case, the variance of the estimated frequency peak is reduced by 37dB if Bayesian inference is used.

B. Limitations in pulse location estimation

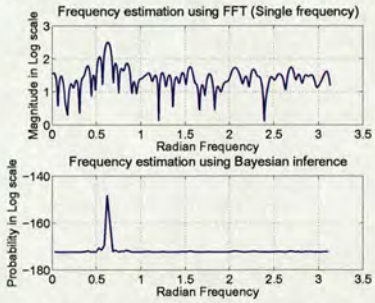
Voice activity detection (VAD) technique is known as an energy detector applied in considerable research areas, and in speech processing in particular [6]–[8]. The method introduces a low-variance spectrum estimation and determines an optimal threshold based on the noise statistics [9]. It only depends on the expected noise power spectral density and the variance of a signal-to-noise ratio (SNR) measure estimated during periods of non-speech activity. The technique calculates the SNR measure and the adaptive threshold for each spectral bin and then compares them to give the final decision of

Yan Yan and James Hopgood are with the Institute for Digital Communications, School of Engineering and Electronics, University of Edinburgh, UK. e-mail: {Y.Yan, James.Hopgood}@ed.ac.uk.

Vassilis Sboros is with Department of Medical Physics, University of Edinburgh, UK. Email: vassilis.sboros@ed.ac.uk.



(a) Simulated signal with single frequency in the time domain.



(b) Probability of the estimated frequency.

Fig. 1. FFT spectrum and Bayesian inference of the simulated signal with single frequency.

whether the signal is the speech or not. From the observation of responses from both MBs and solid spheres, the pulse location estimation of the echo signals in ultrasound imaging is similar to the signal burst detection in speech detection. However, if the SNR of the signal is low or the amplitudes are small, the performance of the VAD is poor.

Hilbert transform (HT) is commonly used for envelope detection in various applications, especially in signal processing for communications [10] and geophysics [11]. It can extract the envelope of the modulated signal accurately [12]. The envelope contains information on energy of the signal [13]. In [14], Verbeek introduced the HT into ultrasound application. However, when the signal is embedded in noise, the HT may fail as it is hard to choose an appropriate threshold and thus can not offer accurate estimation of starting and ending points, which are directly related to pulse locations. Additionally, if only one dominant frequency is present in the signal, the HT can detect the instant frequency with the help of other decomposition techniques but it is not able to estimate multiple frequency components at the same time. As a result, all the abovementioned techniques only operate in time domain. In order to take the frequency information into consideration, a more complicated technique is needed.

There is few information in the ultrasonic literature about the joint estimation of pulse locations and frequencies, which

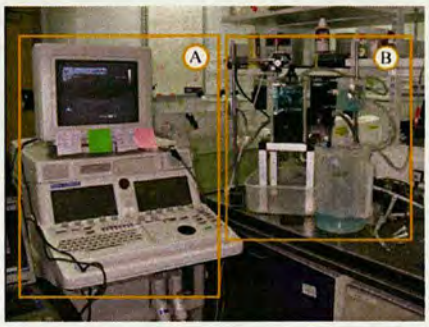


Fig. 2. The equipments used for acquiring ultrasound echo signals.

can be further developed, especially for multiple pulse echo signals from MBs. In the present study, a new estimation system for ultrasound echo signals is proposed. Firstly, the whole procedure is outlined in Section II. The experimental setup is briefly described and then the algorithm and signal model are presented in detail. Secondly, in Section III, simulation results are reported to evaluate the system performance, and the new algorithm for estimation is applied to the measured echo signals from both SCSs and MBs. Finally, the conclusions are in Section IV.

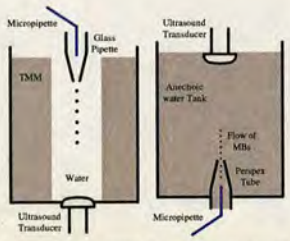
II. Methods

A. Experimental setup

A modified ultrasound transducer (Sonos5500 Philips Medical Systems, Andover, MA, USA) was used to acquire echo signals from both SCSs and MBs. Fig. 2 shows a whole view of the equipments needed in the experimental measurements. The left part of the figure, denoted by A, is the ultrasound transducer, which transmits an ultrasound wave and receives its reflection from the scatterers. The sensitive range of receive frequency is set between 1.2MHz and 4.5MHz. Then the raw echo signals from the scatterers are preamplified, collected and stored in a computer. The right part of the figure, denoted by B, is a set of acquirement facility, which contains a water tank and some perspex tubes.

1) Acquirement of echo signals from SCSs: In order to clearly explain how the acquirement facility works, Fig. 3 shows a simple illustration of experimental setup for acquiring echo signals from SCSs and MBs. As the density of an SCS is larger than water whereas the density of an MB is smaller than water, the equipment setups for them are different, as shown in Fig. 3.

In Fig. 3(a), a beaker was used as a basic tank to measure the scattering from SCSs. A 4cm diameter hole was drilled at the base of the tank and a 25μm thickness Mylar® film was glued to the base to provide an acoustic window. The central cylindrical space defined by the circular acoustic window at the base were filled with degassed water. Other spaces are filled with the tissue mimicking materials, which is used to avoid multiple reflections. A glass pipette was placed at the top of the tank and its bottom tip with 1mm internal diameter was held



(a) Experimental setup for acquiring echo returns from SCSs. (b) Experimental setup for acquiring echo returns from MBs.

Fig. 3. Illustration of experimental setups for acquiring the echoes from SCSs and MBs.

at the center of the tank. Moreover, the measurements were carried out at 7.5cm from the bottom of the tank. The distance is chosen because it was able to put the SCS under a near-plane wave and make positioning errors less critical [15]. Another micropipette was inserted through the larger glass pipette into the tank to ensure the alignment of the beam with the path of SCSs.

A small batch of SCSs with a variety of radii, ranging from 29.5μm to 57.5μm, were held inside the micropipette and then introduced to the tank individually. They descended towards the bottom of the tank with the guidance of the gravity. The ultrasound probe was positioned at the center bottom of the tank to achieve maximum echo amplitude from the tip of the micropipette.

2) Acquirement of echo signals from MBs: The facility for acquiring echo signals from MB is similar to the one used for SCSs. In Fig. 3(b), The tank was filled with anechoic water, which do not produce echoes. A Perspex tube was put at the center bottom of the tank with an 8mm internal diameter. The tip of the glass micropipette, whose diameter is approximately 100μm, was placed at the center of the Perspex tube. Single MBs are released from the micropipette individually. The flow of them were directed towards the face of an ultrasound probe, which was placed at the top of the tank. The distance between the end of the tube and the face of the probe was 7.5cm. The tube and the micropipette were composed of a hydrodynamically focused flow system for isolation of single MBs in a well determined flow path. In this way, the insonation of single MBs can thereby be well calibrated. The stream of MBs was first diluted by introducing filtered water into the infusion path. Then the diluted suspension of MBs was supplied to the micropipette.

B. An estimation system using a parametric signal model

After all responses from MBs and SCSs are obtained from the experimental measurements, the requirement of an estimation system which can extract the information of echo signals in both the time and the frequency domains automatically is of significant importance. An estimation system is proposed in this paper to characterise the different echo signals. According to these difference, the discrimination between the responses

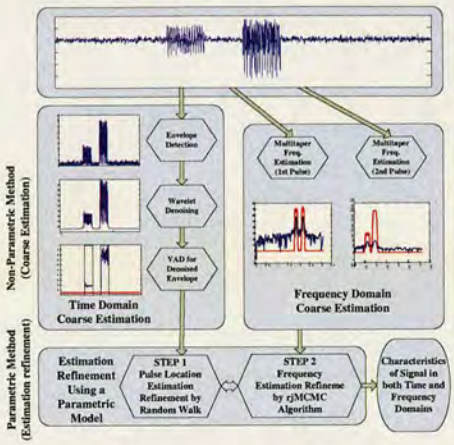


Fig. 4. Diagram for the procedure of the estimation system.

from MBs and SCSs can be exploited. It will offer more insights to the design of transmit pulsing regimes.

The proposed system can be divided into two parts, which are displayed in Fig. 4. The first part is the coarse estimation for pulse locations and frequency components by using *non-parametric* method; the second part is to refine the estimation by using a *parametric* model within a Bayesian framework. Details are described in the following sections.

1) Coarse estimation for pulse locations: As mentioned above, in the case of low SNR, the envelope detected by the Hilbert transform is not of adequate quality. In order to obtain a clearer envelope of the multiple pulse signal, wavelet denoising method can be adopted as it has more advantage than traditional filtering approaches, e.g., it is non-linear and can be applied to non-stationary signals. However, the method may fail when there are closely-spaced pulses present in the measured echo signals. Similarly, the performance of VAD is poor if the SNR of the signal is low or the amplitudes are small.

Although both of the two traditional techniques, the HT and the VAD, used as the energy detector have their own limitations, which can not be neglected in ultrasound imaging, the combination of these two can improve the detection accuracy. Firstly, the HT is applied to the original signal and used for envelope detection. Secondly, the envelope is denoised by stationary wavelet transform, which is selected for its time-invariant property [16]. Finally, the VAD algorithm is performed for the denoised envelope. Following the above procedure, the proposed method can further give better estimation for both closely-spaced pulses and small amplitude pulses. Furthermore, the method can tolerate lower SNR down to 5dB by experimental observation.

2) Coarse estimation for frequency components: Independently, for coarse estimation for frequency, the discrete fourier transform (DFT) is often used to analyze the signal in the frequency domain. Andrieu and Doucet proposed the frequency

estimation algorithm based on the DFT in [17], which may have the problem of overestimating the order of frequencies.

Based on the DFT, there are many modified *non-parametric* methods that provide a better trade-off between sidelobe reduction and the width of spectral peaks. Multitaper spectrum [18] is one of them and can improve the quality of the spectrum a lot. In this technique, several data windows are used on the same original data record to obtain several modified periodograms. Then these periodograms are averaged to produce a multitaper spectrum. By reducing the variance, a much cleaner spectrum is achieved. In addition, the bias and the resolution loss can also be reduced for properly designed tapers.

3) Estimation refinement using a parametric model: The coarse estimation techniques discussed above are all *non-parametric* methods, which are simple but cannot provide acceptable accuracy. In addition, for coarse estimation of frequency, the multitaper spectral estimation only works for single pulse segment signal and cannot provide the number of frequencies automatically. Other model selection schemes need to be incorporated, such as Akaike-information-theoretic (AIC) criterion and Minimum description length (MDL) criterion [19]. For experimental measured echoes from MBs and SCSs, the number of pulse segments in the signal and the number of frequency components in each pulse segment are all unknown. A more advanced estimation technique, using a *parametric* model within a Bayesian framework, can solve these problems and offer more accurate estimates.

As the measured transmit pulse in ultrasound appears to be composed of several cycles of a harmonic signal, the multiple pulse echo signals can be modeled as several segments of sum of sinusoids. Assume there are m pulses in the observed signal with N data points. The multiple pulses model can be defined as follows: ($T_0 \triangleq 1$, $T_{2m+1} \triangleq N$ and $i = 0, \ldots, 2m$)

$$
\begin{aligned}
D_0 \quad &: \quad x(t) = n(t) \\
D_{k_m} \quad &: \quad x(t) = \begin{cases} n(t) & \text{if } T_i < t \leq T_{i+1} - 1,\ i \text{ even} \\ x_i(t) + n(t) & \text{if } T_i < t \leq T_{i+1} - 1,\ i \text{ odd} \end{cases}
\end{aligned}
$$

where $x_i(t) = \sum_{j=1}^{k_i} a_{c_{j,k_i}} \cos(\omega_{j,k_i} t) + a_{s_{j,k_i}} \sin(\omega_{j,k_i} t)$. The signal model can also be written in a vector-matrix form:

$$\mathbf{x} = \mathbf{G}(\boldsymbol{\omega}_{k_m}, \boldsymbol{T}_{2m})\mathbf{a}_k + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is the zero-mean white Gaussian noise with variance σ_k^2 and

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \boldsymbol{G}_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \boldsymbol{G}_m \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The $\mathbf{G}$ matrix has the size of $2N \times \sum_{i=1}^{m} k_i$. It contains the information about changepoints $(T_1, T_2, \ldots, T_{2m})$, which are related to the positions of multiple pulses in the echo signal. Each component $\boldsymbol{G}_j (j = 1, \ldots, m)$ in $\mathbf{G}$ matrix represents a single pulse, which has its own parameters. For each G_j, the number of sinusoids and other parameters $\boldsymbol{\theta}_k \triangleq (\boldsymbol{\omega}_k, \boldsymbol{a}_k, \sigma_k^2)$ are all unknown. As far as each segment is concerned, the $\boldsymbol{G}_j$ matrix can be defined as:

$$\boldsymbol{G}_j = \begin{bmatrix} E(\omega_{k_1} t(T_{2j-1})) & \cdots & E(\omega_{k_m} t(T_{2j-1})) \\ E(\omega_{k_1} t(T_{2j-1} + 1)) & \cdots & E(\omega_{k_m} t(T_{2j-1} + 1)) \\ \vdots & \vdots & \vdots \\ E(\omega_{k_1} t(T_{2j} - 1)) & \cdots & E(\omega_{k_m} t(T_{2j} - 1)) \end{bmatrix}$$

where $E(\cdot) \triangleq [\cos(\cdot), \sin(\cdot)]$. Moreover, T_{2j-1} and T_{2j} are two corresponding changepoints at the start and end points respectively for each pulse segment. For each $\boldsymbol{G}_j$, k_m may have different values, which implies the different number of frequencies and their values in different pulses.

The likelihood function can be easily obtained according to the signal model:

$$p(\mathbf{x}|\{k, \boldsymbol{\theta}_k\}_m, \boldsymbol{T}_{2m}) = (2\pi\sigma_k^2)^{-N/2} \times \exp\left\{-\frac{1}{2\sigma_k^2} \parallel \mathbf{x} - \mathbf{G}(\boldsymbol{\omega}_{k_m}, \boldsymbol{T}_{2m})\boldsymbol{a}_k \parallel^2\right\}, \quad (2$$

where m in $\{k, \boldsymbol{\theta}_k\}_m$ represents different pulse segments and $\|A\|^2 \triangleq A^T \cdot A$.

4) Reversible jump MCMC algorithm for estimation refinement: Fourier transform based methods can only offer the width of spectral peaks, which depends on the duration of the data set. It still needs other methods to determine the accuracy to which the peak frequency can be determined. However, the Bayesian posterior probability for the signal frequency provides this information directly [5] as discussed in the previous section. Moreover, in some situations the noise is not well understood or the data may contain a greater complexity of phenomena than the current signal model incorporates, the use of the DFT is not optimal; more information can be extracted from the data if more sophisticated techniques are used [20]. Bayesian inference can readily handle these situations. The analysis provides much more reliable and informative results when prior knowledge of the shape of the signal and noise statistics are incorporated.

According to Bayesian inference, samples from the posterior distribution can be drawn given the appropriate prior distributions. These priors reflect the degree of belief of the relevant values of the parameters. The joint prior distribution can be considered as several product of independent parameter priors. T_{2m} represents the pulse locations; k_m is the number of frequencies in pulse segment m; $\boldsymbol{\omega}_{k_m}$ is the vector of frequency components in segment m; σ_k^2 denotes the noise variance in the signal, assumed the same in different pulse segments; δ_m^2 and Λ_m are expected SNR and expected number of sinusoids in segment m respectively, which represent hyper-parameters.

TABLE I
PRIOR DISTRIBUTIONS FOR INDEPENDENT PARAMETERS

Parameters	Prior Distributions
T_{2m}	Uniform Distribution
k_m	Truncated Poisson Distribution
ω_{k_m}	Uniform Distribution
σ_k^2	Jeffrey's uninformative prior
δ_m^2	Inverse Gamma Distribution
Λ_m	Gamma Distribution

Following the priors chosen in Table I, the joint prior distribution can be expressed in (3):

$$\begin{aligned} p(\{k,\boldsymbol{\theta}_k\}_m,\boldsymbol{T}_{2m}) &= p(\{k,\boldsymbol{a}_k,\boldsymbol{\omega}_k\}_m|\sigma_k^2)p(\sigma_k^2)p(\boldsymbol{T}_{2m}) \\ &\propto \Big(\frac{\Lambda_m^{k_m}}{k_m!}\exp(-\Lambda_m)\times\frac{1}{|2\pi\sigma_k^2\Sigma_{k_m}|^{1/2}}\times\frac{1}{\pi^{k_m}} \\ &\quad\times\exp[-\frac{\boldsymbol{a}_{k_m}^T\Sigma_{k_m}^{-1}\boldsymbol{a}_{k_m}}{2\sigma_k^2}]\Big)_m \\ &\quad\times\frac{1}{\sigma_k^2}(\frac{1}{N-1}\frac{1}{N-2}\cdots\frac{1}{N-2m}), \end{aligned} \tag{3}$$

where $\boldsymbol{\Sigma}_{k_m}^{-1}=\delta_m^{-2}\mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})$.

The posterior distribution in (4) can be obtained after integrating out the nuisance parameters: amplitudes $\boldsymbol{a}_k$ and noise variance σ_k^2 based on Bayes's rule.

$$\begin{aligned} p(\boldsymbol{T}_{2m},\{k,\boldsymbol{\omega}_k\}_m|\mathbf{x}) &\propto (\gamma_0+\mathbf{x}^T\mathbf{P}_{k_m}\mathbf{x})^{-(N+v_0)/2} \\ &\quad\times\frac{(\Lambda_m/[(\delta_m^2+1)\pi])^{k_m}}{k_m!} \end{aligned} \tag{4}$$

where

$$\mathbf{P}_{k_m} = \mathbf{I}_N-\mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{M}_{k_m}\mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m}) \tag{5}$$

$$\mathbf{M}_{k_m}^{-1} = \mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})+\boldsymbol{\Sigma}_{k_m}^{-1}. \tag{6}$$

Although the posterior distribution is simplified, it is still highly non-linear, which means the closed form of $p(\boldsymbol{T}_{2m},\{k,\boldsymbol{\omega}_k\}_m|\mathbf{x})$ can not be obtained. Therefore, a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm is introduced to sample from the complicated joint posterior distribution and then to estimate the multiple pulse locations and frequency contents for each pulse segment simultaneously.

In standard MCMC algorithm, instead of sampling from the joint posterior distribution directly, an ergodic Markov chain $(\boldsymbol{T}_{2m}^{(i)},\{k^{(i)},\boldsymbol{\omega}_k^{(i)}\}_m)$ whose equilibrium distribution is the specific joint posterior distribution given by (4) is formulated. The simulation is run long enough to reach the stationary distribution. As the model order of the signal is also unknown, the reversible jump technique (rjMCMC) is able to jump between subspaces of different model orders. Then the Maximum A Posterior (MAP) estimator is used to determine the value of parameters that can maximize the conditional probability densities, such as in (7):

$$\begin{aligned} \hat{K}_m &= \arg\max\hat{P}(k_m|\mathbf{x}) \\ \hat{\boldsymbol{\omega}}_{k_m} &= \arg\max\hat{P}(\boldsymbol{\omega}_{k_m}|K=\hat{K}_m,\mathbf{x}) \end{aligned} \tag{7}$$

Based on the initial guesses given by aforementioned combination algorithm of the VAD and the wavelet denoising for envelope detection, a random walk perturbation is adopted as the proposal distribution for refinement of the pulse location estimates. Specifically, the update of each changepoint depends on its previous value and performs a local exploration of the initial guess, which can be described as:

$$T^*|T\sim\mathcal{N}(T,\sigma_T^2). \tag{8}$$

where T and T^* are previous state and new state of the changepoint respectively. $\mathcal{N}(\cdot)$ represents the normal distribution with mean T and variance σ_T^2.

For frequency estimation, in the proposed algorithm, a multitaper spectrum is adopted as the proposal distribution to provide the initial guess for frequency estimation. The rjMCMC algorithm is then used to explore the regions around obvious peaks in the multitaper power spectrum. After the pulse locations are coarsely estimated, frequency estimation is performed for different pulse segments. For each segment m, there are three moves to be selected: *birth* and *death* moves require the dimension changes according to the state of Markov chain whereas the *update* move only refines the frequencies within the same dimension.

birth: propose a new sinusoid randomly on $(0,\ \pi)$.

death: remove an exist sinusoid randomly.

update: Iterations are performed for all k frequencies. With probability λ, (λ is a real number between 0 and 1), an Metropolis-Hasting (MH) step with invariant distribution $p(\omega_{j,k}|\mathbf{x},\boldsymbol{\omega}_{-j,k})$ and proposal distribution $q_1(\boldsymbol{\omega}_{\boldsymbol{j,k}}^*|\boldsymbol{\omega}_{j,k})$ are performed. $\boldsymbol{\omega}_{-j,k}\triangleq(\omega_{1,k},\ldots,\omega_{j-1,k},\omega_{j+1,k},\ldots)$ represents the frequency vector that do not have the current frequency $\omega_{j,k}$; Similarly, with probability $1-\lambda$, the MH step performed with invariant distribution $p(\omega_{j,k}|\mathbf{x},\boldsymbol{\omega}_{-j,k})$ and proposal distribution $q_2(\boldsymbol{\omega}_{\boldsymbol{j,k}}^*|\boldsymbol{\omega}_{j,k})$.

In the new estimation system, the proposed distribution q_1 is based on multitaper spectrum, which is independent of previous state and explore the regions around obvious peaks in the spectrum. Whereas q_2 is simply chosen as random walk perturbation, which is a zero-mean Gaussian random variable with variance σ_{RW}^2, of the previous frequency: $\omega_{j,k}^*|\omega_{j,k}\sim\mathcal{N}(\omega_{j,k},\sigma_{RW}^2)$.

In addition, in each segment m, the probabilities for choosing birth, death and update moves, denoted as b_{k_m}, d_{k_m} and u_{k_m} respectively, are described in [17] in detail.

The refinement of the model parameters, for both pulse locations and frequency components, using a *parametric* model with numerical Bayesian method, consists of two steps in each iteration. First, the pulse locations are updated by random walk perturbation; second, after the estimation of pulse locations, frequency contents for each pulse can be updated using rjMCMC algorithm for a given specific set of changepoints. The procedure is summarized in Algorithm 1.

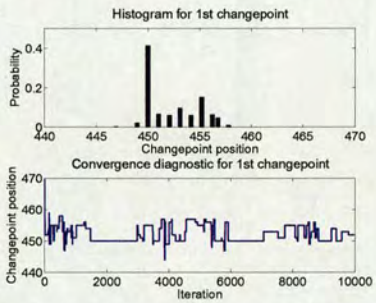
III. RESULTS

A. Performance evaluation on a simulated signal

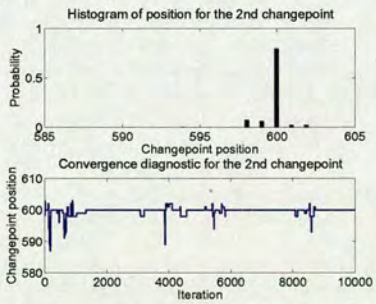
According to the characteristics of the echo signals, the synthetic signal is simulated as several pulse segments with a sum of sinusoids in each segment. This synthetic signal is used to evaluate the performance of the whole estimation system.

Algorithm 1 A rjMCMC algorithm for estimation refinement

Initialization: set $(\{k, \boldsymbol{\theta}_k\}_m^{(0)}, \boldsymbol{T}_{2m}^{(0)})$.
Iteration:
for $i = 1$ to $numIteration$ **do**
    Update each changepoint $\boldsymbol{T}_{2m}^{(i)}$ using random walk.
    For each pulse segment m, update frequency contents:
    a).Sample hyperparameters Λ and δ^2.
    b).Sample u from $U_{(0,1)}$. (*uniform distribution*)
    if $u \leq b_{k_m^{(i)}}$ **then**
        perform *birth move* of a new frequency
    else if $(u \leq b_{k_m^{(i)}} + d_{k_m^{(i)}})$ **then**
        perform *death move* of an existing frequency
    else
        perform *update move* of a frequency randomly
    end if
    c).Sample nuisance parameters $\boldsymbol{a}_{k_m}$ and σ_k^2.
end for



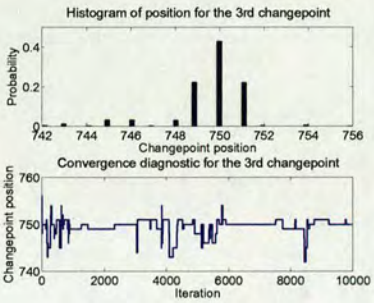
(a) Position estimation refinement for the 1^{st} changepoint.



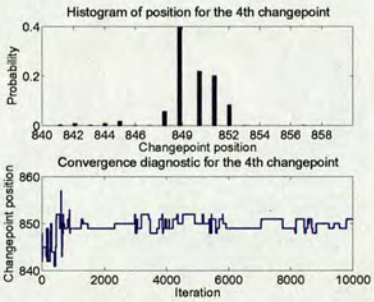
(b) Position estimation refinement for the 2^{nd} changepoint.

Fig. 5. Estimation of positions for the first pulse segment in the simulated signal.

The synthetic signal has 1500 data points, the same length as the experimental echo signal. It consists of two pulses. The first pulse locates between (450, 600) and has frequency components at ($0.6\pi = 1.8849$, $0.7\pi = 2.1991$); the second one locates between (750, 850) and has frequencies at ($0.2\pi = 0.6283$, $0.3\pi = 0.9425$, $0.32\pi = 1.0053$). The whole estimation system is applied to the synthetic data. The number of iterations was set to 10000.



(a) Position estimation refinement for the 3^{rd} changepoint.



(b) Position estimation refinement for the 4^{th} changepoint.

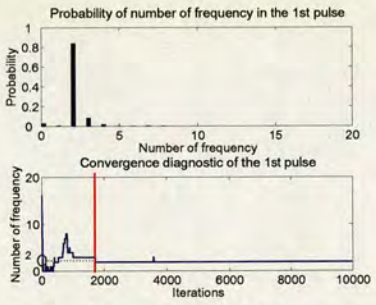
Fig. 6. Estimation of positions for the second pulse segment in simulated signal.

For the pulse locations estimation, if the coarse estimation is not incorporated, a complicated method for sampling from a target distribution is needed and the probability distributions of pulse locations may take a very long time to converge to their true values. In order to reduce the computational complexity, it is always useful to obtain a rough estimate of the location if the target distribution, i.e., a point estimate of the parameters in the model on the basis of some simple, noniterative technique [21]. As a result, the combination of the VAD technique and the HT used as crude estimation for pulse locations can make the convergence much faster.

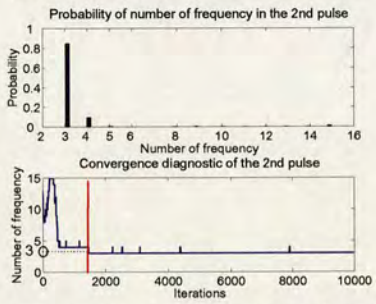
As seen in Fig. 5 and Fig. 6, after 1000 iterations, all the changepoints, related to the starting and ending points of pulses, reach their stationary distributions. The results of the coarse estimation and the estimation refinement for pulse locations are compared in Table II. The error is calculated as the sum of difference between true values and estimated values of each pulse location. The error after estimation refinement is much less than that of coarse estimation only.

TABLE II
COMPARISON BETWEEN THE COARSE ESTIMATION AND THE ESTIMATION REFINEMENT OF PULSE LOCATIONS (START & END POINTS) SNR = 5DB

True Locations	450	600	750	850
Coarse Estimation	460	602	760	850
$\lvert\triangle\epsilon\rvert$	1	4	2	1
Estimation Refinement	452	600	750	850
$\lvert\triangle\epsilon\rvert$	2	0	0	1



(a) Number of frequency estimation for the 1^{st} pulse.

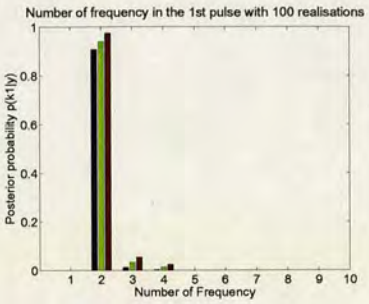


(b) Number of frequency estimation for the 2^{nd} pulse.

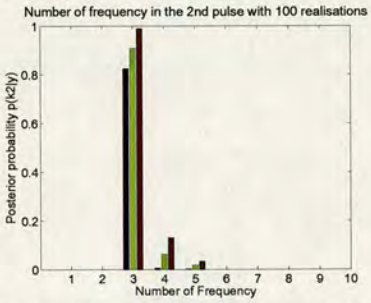
Fig. 7. Number of frequency estimation and the convergence diagnostics.

For the 1^{st} and 2^{nd} pulses, the probability of number of frequencies and the convergence diagnostics are shown in Fig. 7(a) and Fig. 7(b). It can be seen clearly that after about 3000 iterations, the probability of number of frequencies for the 1^{st} pulse converges and similarly, the probability of number of frequencies for the 2^{nd} pulse converges after about 1500 iterations. As a result of these convergence, for the 1^{st} pulse, the first 3000 iterations are discarded as the burn-in period. Then the histogram is plotted by using the rest of samples; for the 2^{nd} pulse, the histogram is plotted by discarding the first 1500 iterations.

Furthermore, the above experiments has been repeated 100 times with different noise realizations and amplitudes and phase components. On average, the error of coarse estimation for pulse locations is 20 data points whereas the error after estimation refinement is 5 data points. For frequency estimation,



(a) Histogram showing the mean, $mean - SD$, $mean + SD$ of the estimate of $p(k_1|x_1)$ for the first pulse.

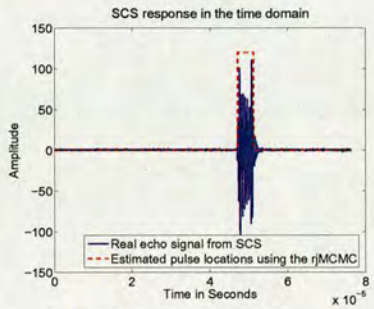


(b) Histogram showing the mean, $mean - SD$, $mean + SD$ of the estimate of $p(k_2|x_2)$ for the second pulse.

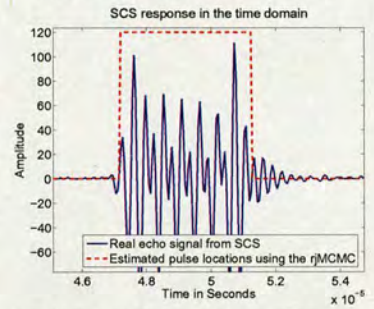
Fig. 8. Histograms of number of frequencies using 100 noise realizations.

Fig. 8(a) and Fig. 8(b) show the histogram of estimates for 1^{st} and 2^{nd} pulses from 100 realisations. In Fig. 8, the mean and the SD value of the estimated posterior probability of the number of frequencies are illustrated. The green bars denote the mean values of the estimated number of frequencies; the blue and red bars denote the $mean - SD$ and $mean + SD$ of the estimates respectively. The peaks in all three statistics at $k_1 = 2$ in Fig. 8(a) and $k_2 = 3$ in Fig. 8(b) indicate the good estimation performance of the developed algorithm.

Moreover, the estimated frequency values obtained from 100 Monte Carlo runs for two pulse segments of the simulated signal are compared to the ground truths, and are also compared to the results using the *non-parametric* estimation method. Table III(a) provides the comparison for the first pulse segment. Compared with the ground truths, both the *non-parametric* method and the *parametric* method provide good mean estimates. However, in terms of the standard deviations (SDs), the *parametric* method outperforms the *non-parametric* method with much smaller variations. Table III(b) gives the comparison for the second pulse segment. The *parametric* method not only provides more accurate mean estimates and smaller SDs, but can distinguish the two closely-spaced frequencies as well. In contrast, the *non-parametric*



(a) Actual SCS response in the time domain.



(b) Enlarged actual SCS response in the time domain.

Fig. 9. Display of the actual experimental SCS response in the time domain.

method regard the two closely-spaced frequencies as one single frequency, which signifies its poor performance of frequency resolution.

TABLE III
COMPARISON OF FREQUENCY VALUES BETWEEN THE *non-parametric* METHOD AND *parametric* METHOD FOR TWO PULSE SEGMENTS.

Freqs.	Ground Truth	*non-parametric* method (mean ± SD)	*parametric* method (mean ± SD)
1^{st}	2.1991	2.2000 ± 0.0364	$2.1995 \pm 4.3051 \times 10^{(-4)}$
2^{nd}	1.8849	1.8870 ± 0.0530	$1.8850 \pm 7.2931 \times 10^{(-4)}$

(a) Frequency value comparison for the 1^{st} pulse segment

Freqs.	Ground Truth	*non-parametric* method (mean ± SD)	*parametric* method (mean ± SD)
1^{st}	0.6283	0.6200 ± 0.2097	$0.6285 \pm 4.3019 \times 10^{(-4)}$
2^{nd}	0.9425	0.9790 ± 0.2196	$0.9387 \pm 4.8732 \times 10^{(-4)}$
3^{rd}	1.0053	0.0	$1.0027 \pm 5.3047 \times 10^{(-4)}$

(b) Frequency value comparison for the 2^{nd} pulse segment

B. Estimation for measured ultrasound echo signals

1) Estimation of an SCS response: Fig. 9(a) displays a typical SCS response, which has a single pulse segment. This typical response is first modeled as a single pulse signal, which is composed of several sinusoids. The proposed estimation system is then carried out and repeated 100 times to the SCS response. Finally, these results are averaged and calculated to give the means and the SD values of the estimated parameters of interest, i.e. pulse locations, number of frequencies and the corresponding frequency values.

TABLE IV
ESTIMATED PULSE LOCATION FOR THE SCS RESPONSE

change-points	1^{st}	2^{nd}
Means	927	1007
SD values	1	4

TABLE V
SPREAD OF MAP ESTIMATES OF NUMBER OF FREQUENCY FOR THE SCS RESPONSE

K	MAP
$0 \sim 7$	0
8	1
9	61
10	35
11	3
$\geqslant 12$	0

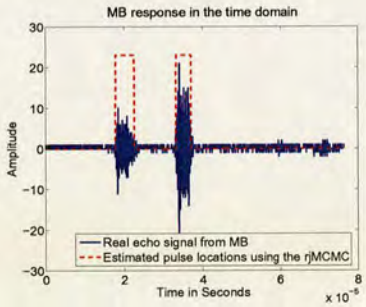
TABLE VI
ESTIMATED FREQUENCY VALUES FOR THE SCS RESPONSE

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	1.8138	0.0734
2^{nd}	2.1687	0.2858
3^{rd}	2.4339	0.3511
4^{th}	2.6924	0.3516
5^{th}	2.9891	0.4241
6^{th}	3.2320	0.3157
7^{th}	3.3145	0.2756
8^{th}	3.3941	0.4564
9^{th}	3.5345	0.6786

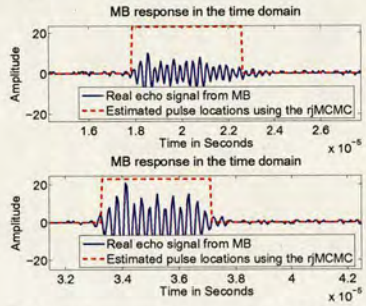
The estimated pulse location in the time domain is presented in Table IV with the means and the SD values. Moreover, the estimates are also marked in Fig. 9(a) using a red dotted line to denote the start and end points of the pulse segment. Fig. 9(b) shows an enlarged version of Fig. 9(a) with explicit indication of the estimated pulse location.

Furthermore, for each run of the estimation algorithm, the standard Bayesian estimate of the model order for a particular pulse segment i, i.e. the number of frequency, is the MAP value of K_i, which is the value of k_i that maximizes $p(k_i|\mathbf{x}_i)$. In this measured SCS response, there is only one pulse segment. The MAP value of K in the single pulse can be obtained in each run of the algorithm. Then with 100 replications of the algorithm, the spread of MAP values are shown in Table V. It can be seen that the most probable number of frequency is 9, which appears 61 times in the 100 replications. Moreover, the estimated frequency values in the single pulse segment are provided in Table VI with the means and the SD values.

2) Estimation of an MB response: Unlike SCS responses, most of MB responses are composed of multiple pulse segments. They can be modeled as a multiple pulse signal model, each of the pulse is modeled as a sum of sinusoids. One typical example with two pulse segments is shown in Fig. 10(a).



(a) Measured MB response in the time domain.



(b) Enlarged measured MB response in the time domain.

Fig. 10. Display of the actual experimental MB response in the time domain.

It is chosen because it not only can describe the procedure of estimation for multiple pulses clearly, but also will not make the computations too complicated using the *parametric* estimation. Therefore, it is a good compromise between an explicit demonstration of the developed estimation algorithm and the computational complexity that a typical example of ultrasound responses requires.

The developed estimation system is also repeatedly applied to the MB response with 100 times. The estimation of the number of pulses, and the corresponding pulse locations are given in Table VII. The estimated start and end points of two pulse segments are also marked as red dotted lines in Fig. 10(a). The enlarged version for each single pulse segment is illustrated in Fig. 10(b) for clarity. Moreover, the spread of MAP values of the number of frequencies for two pulse segments are shown in Table VIII. The result in Table VIII(a) indicates that 63% of the time, the detected number of frequency is 4 for the first pulse. For the second pulse segment, the most probable number of frequency is 5, which appears 58 times in 100 replications of the estimation. Furthermore, the estimated frequency values of the first pulse are presented in Table IX(a) and those of the second pulse are presented in Table IX(b). These resulting estimates are given with both mean and SD values.

TABLE VII
ESTIMATED PULSE LOCATION FOR THE MB RESPONSE

Change-point	1^{st}	2^{nd}	3^{rd}	4^{th}
Means	351	444	654	730
SD values	2	4	2	2

TABLE VIII
SPREAD OF MAP ESTIMATES OF NUMBER OF FREQUENCY FOR TWO PULSE SEGMENTS IN THE MB RESPONSE

K_1	MAP
$0 \sim 2$	0
3	10
4	63
5	13
6	9
7	3
8	2
$\geqslant 9$	0

(a) Spread of MAP estimates of number of frequency for the 1^{st} pulse

K_2	MAP
$0 \sim 2$	0
3	6
4	27
5	58
6	6
7	3
$\geqslant 8$	0

(b) Spread of MAP estimates of number of frequency for the 2^{nd} pulse

TABLE IX
ESTIMATED FREQUENCY VALUES OF TWO PULSES IN THE MB RESPONSE

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	2.4390	0.4282
2^{nd}	2.7667	0.4597
3^{rd}	3.1077	0.5327
4^{th}	3.4059	0.4632

(a) Estimated frequency values for the 1^{st} pulse

Estimates(Freq.)	Means(MHz)	SD values(MHz)
1^{st}	2.0710	0.4228
2^{nd}	2.6467	0.4207
3^{rd}	2.9594	0.4414
4^{th}	3.3252	0.6160
5^{th}	3.3841	0.6210

(b) Estimated frequency values for the 2^{nd} pulse

C. Discussions

The whole algorithm has been carried out for all the available data sets. The investigation indicates that the proposed estimation system can estimate the number of pulses in the signal and their positions, as well as the number of frequencies in each pulse and their corresponding values with very small errors simultaneously. In all, the results obtained by the new estimation system are all consistent with the theory. Therefore, it can be seen as a robust tool to estimate the information both in the time and frequency domains simultaneously. These attributes detected by the proposed estimation system can lead to an extension to differentiating the echo signals from MBs and tissues, which is not within the scope of this paper but is a prosperous field for MB behavior analysis.

Despite the advantages of this newly developed estimation system in both the time and the frequency domains in a *parametric* way, the limitations of it are twofold. First, in terms of the estimation algorithm itself, the limitation lies in the assumptions made in estimating number of pulses in the signal. The multiple pulses in the echo signal are assumed to be well separated and each of the pulse has a

minimum length. Specifically, if two pulses are closely-spaced, they will be regarded as one pulse; if the width of a single pulse is less than a certain value, it will be regarded as noise. Moreover, the proposed model is demonstrated to fit the measured ultrasound echo signals well rather than being justified in a strict sense, which needs more validation in the future work. Second, in accordance with the requirement of ultrasound imaging techniques, due to the influence of ultrasound transducer in the experimental measurements, the estimated spectral contents may not originate from the true behavior of ultrasound scatterers.

IV. Conclusions

This paper proposes a novel estimation system for echo signals from ultrasound contrast MBs and SCSs. The system introduces a *parametric* model and estimates the parameters within a Bayesian framework. As the posterior density function can not be solved in a closed form and the dimension of the parameters changes, a reversible jump MCMC algorithm is considered and adopted to give more accurate estimation automatically. Moreover, in order to speed up the convergence and reduce the computation cost, the coarse estimation for both time and frequency domains is incorporated.

The advantage of this estimation system is that it allows an automatic estimation of frequencies for each pulse and the pulse locations at the same time. Moreover, it exhibits improved frequency resolution compared to the Fourier analysis based techniques. In addition, the *parametric* model introduced in the paper optimises the estimation of all parameters of interest. In all, the results from estimation system are consistent with the theoretical calculations and deductions.

As a result, the new estimation system reveals more characteristics in both time and frequency domains, which may broaden the research field in ultrasound contrast agents, especially in design of transmit pulsing regimes.

Acknowledgment

The authors would like to thank Biological inspired acoustic systems (BIAS) project to provide the financial support.

References

[1] R. Gramiak, "Echocardiography of the aortic root," *Invest Radio*, vol. 3, pp. 356–366, 1968.

[2] E. Stride and N. Saffari, "Microbubble ultrasound contrast agents, a review," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 217, 2003.

[3] N. de Jong, R. Cornet, and C. T. Lancee, "Higher harmonics of vibrating gas filled microspheres, part one:simulations," *Ultrasonics*, vol. 32, pp. 447–453, 1994.

[4] M. Akay, *Detection and estimation methods for biomedical signals*. San Diego-California-USA: Academic Press, 1999.

[5] P. Gregory, "A bayesian revolution in spectral analysis," in *Bayesian inference and maximum entropy methods in science and engineering*, 2001.

[6] J.-H. Chang, N. S. Kim, and S. K.Mitra, "Vioce activity detection based on multiple statistical models," *IEEE Trans. Signal Process.*, 2006.

[7] Y. D. Cho and A. Kondoz, "Analysis and improvement of a statistical model-based voice activity detector," *IEEE Signal processing letters*, vol. 8, 2001.

[8] J. Sohn, N. S. Kim, and W. Sung, "A statistical model-based voice activity detection," *IEEE Signal processing letters*, vol. 6, 1999.

[9] A. Davis, S. Nordholm, and R. Togneri, "Statistical voice activity detection using low-variance spectrum estimation and an adaptive threshold," *IEEE Trans. Speech Audio Process.*, vol. 14, pp. 412–424, 2006.

[10] Alencar, M. S., and R. J. V. C., *Communication systems*, 2005.

[11] N.Sundararajan and Y.Srinivas, "A modified hilbert transform and its application to self potential interpretation," *Journal of applied geophysics*, vol. 36, pp. 137–143, 1996.

[12] S. L. Hahn, *Hilbert transforms in signal processing*. Artech House Signal Processing Library, 1996.

[13] M. Feldman, "Theoretical analysis and comparison of the hilbert transform decomposition methods," *Mechanical Systems and Signal Processing*, vol. 22, pp. 509–519, 2008.

[14] A. Xander, A. M. Verbeek, L. A. F. Ledoux, P. J. Brands, and A. P. G. Hoeks, "Baseband velocity estimation for second-harmonic signals exploiting the invariance of the doppler equation," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 1217–1226, 1998.

[15] V. Sboros, S. D.Pye, C. A.MacDonald, J. Gomatam, C. M.Moran, and W. N.McDicken, "Absolute measurement of ultrasonic backscatter from single microbubbles," *Ultrasound in Med.&Biol.*, vol. 31, pp. 1063–1072, 2005.

[16] X. Lu, R. Liu, J. Liu, and S. Liang, "Removal of noise by wavelet method to generate high quality temporal data of terrestrial modis products," *Photogrammatric Engineering and Remote Sensing*, vol. 73, 2007.

[17] C.Andrieu and A.Doucet, "Joint bayesian model selection and estimation of noisy sinusoids via reversible jump mcmc," *IEEE Trans. Signal Process.*, vol. 47, pp. 2667–2676, 1999.

[18] D. Manolakis, V. Ingle, and S. Kogon, *Statistical and adaptive signal processing: Spectral estimation, signal modelling, adaptive filtering and array processing*. McGrawHill, 2003.

[19] P. Stoica and Y. Selen, "Model-order selection: A reivew of information criterion rules," *IEEE Signal Processing Magazine*, pp. 36–47, 2004.

[20] G. Bretthorst, *Bayesian spectrum analysis and parameters estimation*. Springer-Verlag, N.Y., 1988.

[21] A. Gelman, J. B. Carlin, H. S.Stern, and D. B.Rubin, *Bayesian data analysis (Texts in Statistical science)*. Chapman&Hall/CRC, 2004.

Statistical analysis of echo signals from ultrasound contrast microbubbles

Yan Yan [a] James R. Hopgood [a] Vassilis Sboros [b]

[a] *AGB Building, King's Buildings, School of Engineering, University of Edinburgh, Edinburgh, U.K., EH9 3JL*

[b] *47 Little France Crescent, School of Clinical Sciences and Community Health, University of Edinburgh, Edinburgh, U.K., EH16, 4TJ*

Corresponding Author: Yan Yan
Room 2.12, AGB Building, King's Buildings, University of Edinburgh
Edinburgh, U.K., EH9 3JL
Email: Y.Yan@ed.ac.uk
Telephone: +44 131 6505659

Abstract

Nowadays, since the introduction of UCAs to diagnostic Ultrasound contrast imaging techniques, many investigations have focused on theoretical prediction, which are inadequate and questionable. In order to characterise the measured ultrasound echo signals based on *in-vitro* experiments, the influence of ultrasound transducer cannot be neglected. However, general signal processing approaches cannot meet the requirements. This paper develops a *parametric* estimation method based on a proposed signal model, which incorporates the ultrasound transducer characteristics into the signal model as model parameters. Using this newly developed estimation approach, the true spectral content of ultrasound scatterers can be revealed. This will also broaden the research area of the analysis of MB behavior.

Key words: Ultrasound contrast microbubbles, spectral analysid, MCMC algorithms, transducer characteristics

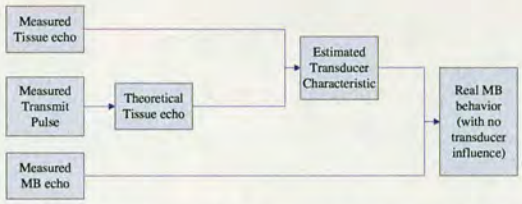


Figure 1. The procedure of extraction of the real MB behavior.

Introduction

Nowadays, the ultrasound echo signals under analysis are based on experimental measurements. In actual measurements, unfortunately, it is possible to calibrate the transmitted acoustic field using a membrane hydrophone, yet it is much more complex to do so for a received beam (Sboros et al, 2005). It is because that the exact centering of microbubbles (MBs) in the transmitted field is very difficult, and it cannot be done with a precision less than 1mm from the central line, which is very important in charactersing the field in every possible direction. As a result, an ultrasound transducer is used to obtain MB responses.

Nevertheless, the transducer has a band limited property. It will have an influence on the estimation of spectral contents of scatterers, which may lead to inaccurate understanding of their behavior, especially for MBs. Previous investigations in the ultrasound literature usually use the transducer spectral sensitivity to normalize the received signals (Krishna et al, 1999) or use a so-called perfect reflector as a means of normalizing backscattered signals (de Jong et al, 1992; Frinking and de Jong, 1998; Marsh et al, 1998). However, in the first method, the transducer spectral information cannot be easily obtained; the second method is not suited to the nonlinear scatterers, e.g. MBs.

In order to take the transducer characteristics into consideration, the diagram in Figure 1 illustrates an intuitive way to recover the true spectra of MBs and thus reveal real MB behavior. Specifically, the characteristics of the transducer, especially its spectrum, are first examined by comparing the theoretical and the measured echo signals from linear scatterers. Then by combining the received experimental measurements of MB echoes with the calibrated transducer characteristics, real MB behavior can be revealed. In this process, how to combine the transducer characteristics and the measured MB responses is a critical issue. It requires many advanced signal processing techniques, which can provide powerful algorithms for inference, estimation and detection.

This paper will dedicate to a proposed signal model taking account of the transducer characteristics as model parameters. Moveover, a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm is developed to estimate the proposed model parameters. The detailed procedure of spectral estimation algorithm will be described in Section . After the evaluation over a simulated signal, the developed estimation algorithm based on the proposed signal model, is applied to measured ultrasound echo signals, which can reveal the true spectra of ultrasound scatterers. The results will be presented in Section 1. Finally, Section 1 will address conclusions.

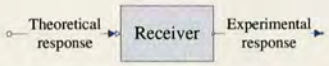


Figure 2. A simple illustration of the receiver calibration.

Method

Calibration of ultrasound transducer

The calibration of the ultrasound transducer characteristics contains two parts: one is the calibration of the magnitude response and the other is the calibration of the phase response. They are achieved using two different approaches and are investigated in the following respectively.

Magnitude response of the ultrasound transducer

A simple illustration of the receiver calibration is depicted in Figure 2. The receiver response can be estimated by looking at the signals at the input and output of the transducer. The magnitude response of the receiver can be calculated by taking the ratio of the energy density of the experimentally measured responses from the SCSs (Sboros et al, 2005) and the corresponding theoretical responses (Hilgenfeldt et al, 2000). Moreover, since the transmit pulse is a narrow-band signal containing a dominant fundamental and a second harmonic spectral component, the ratio of the magnitude responses will only cover a limited bandwidth in the spectrum. To estimate over the full range of the receiver spectral response, it is necessary to use transmit pulses with different incident frequencies and filter them into the fundamental and the second harmonic spectral components. Therefore, the frequency response of the transducer can be estimated by filtering the transmit pulse into fundamental and second harmonic components, calculating the theoretical echo returns at the transducer input, filtering the measured transducer output in the same way, and then calculating the ratio of the energy density of the output to the input.

To be more specific, for each transmit frequency, energy densities of the filtered fundamental and second harmonic components should be calculated separately. Concerning a calibration point, $G_{calibrate_{mag}}$, in the frequency spectrum of the ultrasound transducer based on a single frequency, f_{mag}, it can be calculated in the time domain according to Parseval's theorem:

$$G^2_{calibrate_{mag}}(f_{mag}) = \frac{\text{Energy Density}_{exp}}{\text{Energy Density}_{theory}} = \frac{\int_{t_{min}^{(exp)}}^{t_{max}^{(exp)}} |P_{exp}(t)|^2 dt}{\int_{t_{min}^{(theory)}}^{t_{min}^{(theory)}} |P_{theory}(t)|^2 dt}, \quad (1)$$

where $P_{exp}(t)$ and $P_{theory}(t)$ represent the measured and the theoretical backscattered radiation in the time domain. The time intervals, i.e. $[t_{min}^{(theory)}, t_{max}^{(theory)}]$ and $[t_{min}^{(exp)}, t_{max}^{(exp)}]$, over which the integration are evaluated should contain sufficient durations to include the complete pulse segment for both the theoretical response and the measured response.

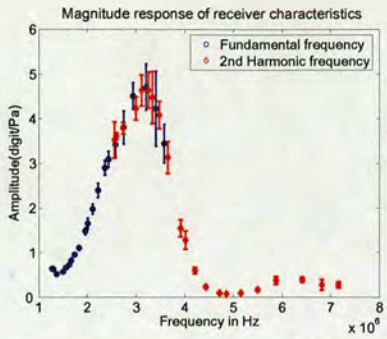


Figure 3. Receiver calibration in the whole frequency range.

A detailed explanation of how to calculate a calibration point in the frequency spectrum of the ultrasound receiver at a specific transmit frequency is presented in (Sboros et al, 2005).

Furthermore, there are 21 different transmit frequencies ranging from 1.2MHz to 4.0MHz. Each of them will be filtered into the fundamental and the second harmonic components and thus there will be 42 calibration points in the spectrum of the ultrasound receiver. Under each transmit frequency, there are several trials of experimental measurements for SCS responses. These SCSs have various radii. With these different trials of measurements, the means and the SD values of the calibration points can be obtained, which are plotted in Figure 3. The circles and diamonds denote the means, and the error bars denote the SD values of the calibrated magnitude gains of the receiver. Moreover, the blue circles and error bars represent results from the fundamental components; the red diamonds and error bars represent results from the second harmonic components. By checking the overlap of the blue circles and the red diamonds, the accuracy of the experimental measurements can be demonstrated and the shape of the receiver bandwidth can be created. In the full range of the transmit frequencies, the SD values are relatively high when the frequencies are between 2.5MHz and 3.5MHz.

In Figure 3, there are only finite frequency values rather than a continuous curve available. In order to incorporate this information into a signal model, these finite number of the magnitude gains against frequencies in the spectrum require interpolation. For simplicity, a polynomial function is assumed so as to fit a curve to the finite numbers in the calibrated magnitude response of the ultrasound receiver. Traditionally, the most common approach to evaluate the best fit is the least squares method Manolakis et al (2000). An alternative is Bayesian linear regression and the derivations can be found in Ruanaidh and Fitzgerald (1996).

When fitting a suitable curve to the finite numbers in the spectrum, whether the estimation approach is based on the least squares method or based on Bayesian linear regression, determining a model order for the polynomial function is required in the first place. Some simple and commonly used model order selection criteria, such as AIC and MDL, have been demonstrated with good performances in solving simple problems (Haykin, 2002). The estimated results of the polynomial coefficients are displayed in Table 1 with means and SD values. The means of the estimated coefficients are used to create a continuous curve for the magnitude response of the receiver, as displayed in Figure 4. It shows a

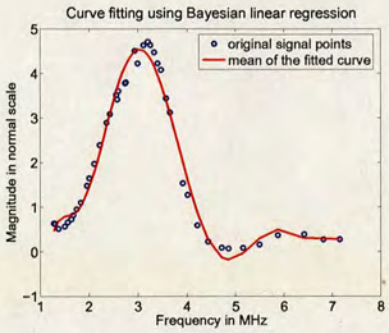


Figure 4. Curve fitting for individual experimental data with known model order.

Table 1
Estimated polynomial coefficients of the magnitude response of the receiver.

Coefficients $\hat{\beta}$	Means	SD values
$\hat{c}_0$	-180.6035	9.3645
$\hat{c}_1$	505.7294	24.4812
$\hat{c}_2$	-578.6615	26.5858
$\hat{c}_3$	352.5759	15.6980
$\hat{c}_4$	-124.9146	5.5281
$\hat{c}_5$	26.5174	1.1927
$\hat{c}_6$	-3.3228	0.1545
$\hat{c}_7$	0.2266	0.0110
$\hat{c}_8$	-0.0065	3.3221×10^{-4}

good match between the finite number of blue circles and the red continuous line.

Phase response of the ultrasound transducer

By calculating the ratio of energy density of the measured ultrasound echo signals to the theoretical echo signals, only the magnitude response can be obtained. However, the magnitude response of the receiver cannot be used to fully characterise the ultrasound receiver. Estimation of the phase response of the receiver is also necessary. There is a different way to estimate the phase response, which is calibrated by taking the ratio of the Fourier transform of the measured responses to the theoretical responses. It is often termed as a transfer function approach. As the Fourier transform of a signal covers the whole range of frequencies, from 0MHz to $fs/2 \approx 9.8$MHz, the transmit pulse which is used to generate the theoretical response, and the measured response do not need to be filtered to the fundamental and the second harmonic separately. The procedure of estimating the phase response based on the Receiver-Transfer-Function (RTF) using a single measured response and a single theoretical response from SCSs can be summarised as:

(1) Calculate the phase response of an experimentally measured echo signal based on

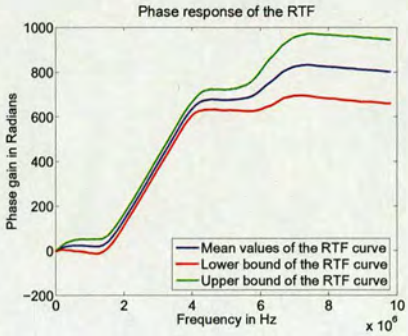


Figure 5. Averaged phase response obtained based on the RTF with 21 different transmit frequencies.

the Fourier transform.

(2) Calculate the phase response of a theoretical echo signal based on the Fourier transform.

(3) Calculate the the phase response RTF by calculating the difference between the phase response of the output and that of the input to the ultrasound receiver according to:

$$G_{calibrate_{phase}} = \arg[H(\omega)] = \arg[F_{output}(\omega)] - \arg[F_{input}(\omega)]. \quad (2)$$

As aforementioned, under each transmit frequency, there are several trials of measured SCSs responses with various radii and a single theoretical SCS response available. For example, there are 9 measured SCS responses with a transmit frequency at 1.67MHz. Therefore, there will be 9 phase responses, which can produce a mean RTF of the phase response. Moreover, considering 21 different transmit frequencies, there will be 21 RTFs of the phase response. Average these phase responses to produce a overall mean RTF of the phase response and the corresponding SDs. Figure 2(a) displays an example with the obtained phase response under the transmit frequency at 1.67MHz. Figure 2(b) depicts the overall RTF of the obtained phase responses and the corresponding SDs.

A proposed signal model

The proposed signal model is in a matrix-vector form as:

$$\begin{aligned} \mathbf{y} &= \mathbf{HS} + \mathbf{n} \\ &= \mathbf{HDa}_k + \mathbf{n}, \end{aligned} \quad (3)$$

where $\mathbf{D}$ contains frequency information $\boldsymbol{\omega}_k$ and $\mathbf{a}_k$ contains the amplitudes for each frequency. Moreover, $\mathbf{n}$ is a zero-mean additive Gaussian noise. The only difference lies in the receiver matrix $\mathbf{H}$, which contains the impulse response of ultrasound receiver in the time domain. The matrix $\mathbf{H}$ is a convolution matrix [1], which is defined as:

[1] A convolution matrix is a matrix formed from a vector, which is able to convert the convolution operation of two vectors to the multiplication operation of two in the time domain.

$$\mathbf{H} \triangleq \begin{bmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \dots & \vdots \\ \vdots & h(1) & \dots & 0 \\ h(M-1) & \vdots & \dots & h(0) \\ \vdots & h(M-1) & \dots & h(1) \\ 0 & \vdots & \dots & \vdots \\ 0 & 0 & \dots & h(M-1) \end{bmatrix}, \tag{4}$$

where

$$\mathbf{h} = [h(0), \quad h(1), \quad \dots, \quad h(M-1)]^T \tag{5}$$

is the impulse response of the calibrated ultrasound receiver in the time domain. It can be easily achieved by being transformed from the frequency spectrum of the receiver $\tilde{\mathbf{H}}$ in the frequency domain. As the magnitude and the phase responses are both assigned a polynomial model as a function of a frequency bin, $\tilde{\mathbf{H}}$ in the frequency domain can be obtained with coefficient vectors $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{H}}(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p) = \left(\sum_{r=0}^{P_1} \theta_m(r) \boldsymbol{\omega}^r \right) \cdot \exp \left[j \left(\sum_{s=0}^{P_2} \theta_p(s) \boldsymbol{\omega}^s \right) \right], \tag{6}$$

where $\boldsymbol{\omega} \triangleq [\omega(0), \quad \omega(1), \quad \dots, \quad \omega(M-1)]^T$ represents the vector of M frequency bins. $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$ are coefficients of the polynomial functions for the magnitude response and the phase response of the receiver respectively. They are assumed to have model order P_1 and P_2.

It is assumed that the noise vector $\mathbf{n}$ is an identically independently distributed Gaussian noise with a variance of σ^2, and all the parameters describing the received echo signal are stationary throughout the entire observation. The likelihood function of the signal model in (3) can be expressed as:

$$p(\mathbf{y}|k, \boldsymbol{\Phi}) = (2\pi\sigma^2)^{-L/2} \cdot \exp \left\{ -\frac{1}{2\sigma^2} [\mathbf{y} - \mathbf{H}\mathbf{D}\mathbf{a}_k]^T [\mathbf{y} - \mathbf{H}\mathbf{D}\mathbf{a}_k] \right\}, \tag{7}$$

where $L = M + N - 1$. N is the length of original echo signal and M is the length of the receiver in the time domain. k is the number of sinusoids in the signal model. Other parameters are denoted as: $\boldsymbol{\Phi} \triangleq (\boldsymbol{\omega}_k, \boldsymbol{a}_k, \sigma^2, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p)$.

Prior distributions for model parameters

The joint prior distribution of all these parameters can be expressed as follows according to their hierarchical structure of the signal model:

$$p(k, \boldsymbol{\Phi}) = p(k|\Lambda)p(\Lambda)p(\boldsymbol{\omega}_k|k)p(\boldsymbol{a}_k|\sigma^2, \delta^2)p(\sigma^2|v_0, \gamma_0)p(\delta^2)p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p). \tag{8}$$

Table 2
Priors for all parameters

Parameters	Prior Distributions
$k\|\Lambda$	Truncated Poisson Distribution
Λ	Gamma Distribution
$\boldsymbol{\omega}_k\|k$	Uniform Distribution $(0,\pi)^k$
$\mathbf{a}_k\|\sigma^2,\delta^2$	Multivariate Normal Distribution
$\sigma^2\|v_0,\gamma_0$	Jeffrey's uninformative prior
δ^2	Inverse Gamma Distribution
$\boldsymbol{\theta}_{m(1:P_1)}$	Gaussian distribution
$\boldsymbol{\theta}_{p(1:P_2)}$	Gaussian distribution

Table 2 displays the prior distributions for the parameters $\boldsymbol{\Phi}$. Most of them are uninformative priors or conjugate priors, as those chosen in (Andrieu and Doucet, 1999). These priors can be multiplied to produce a joint prior distribution directly as they are independent with each other. If $0 < k \leqslant k_{max}$, the joint prior is given in (9); if $k = 0$, the joint prior is stated in (10).

The priors for coefficients of the magnitude and the phase functions, $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$, are independent of other parameters. As there are some rough idea of coefficient values of the magnitude and the phase responses in Section , in which case, a Gaussian distribution centered about a single value is a good choice of the prior for this particular coefficient parameter. In order to simplify the expression of the joint prior distribution and the joint posterior distribution in the following expressions, $p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p)$ will remain in the intact form.

$$
\begin{aligned}
p(k,\boldsymbol{\Phi}|0<k\leqslant k_{max}) \propto & \frac{\Lambda^{k+\epsilon_1-1/2}}{k!}\exp[-\Lambda(1+\epsilon_2)]\cdot\frac{1}{|2\pi\sigma^2\Sigma_k|^{1/2}}\cdot\exp\left[-\frac{\boldsymbol{a}_k^T\Sigma_k^{-1}\boldsymbol{a}_k}{2\sigma^2}\right] \\
& \times\frac{\mathbb{I}_\Omega(k,\boldsymbol{\omega})}{\pi^k}\cdot\frac{\exp[-\gamma_0/(2\sigma^2)]}{\sigma^{2(v_0/2+1)}}\cdot(\delta^2)^{-\alpha_{\delta^2}-1}\cdot\exp\left(-\frac{\beta_{\delta^2}}{\delta^2}\right) \\
& \times p(\boldsymbol{\theta}_m,\boldsymbol{\theta}_p), \qquad (9) \\
p(\boldsymbol{\Phi}|k=0) \propto & \exp\left(-\frac{\gamma_0}{2\sigma^2}\right)\cdot\frac{1}{\sigma^{2(v_0/2+1)}}\cdot e^{-\Lambda(1+\epsilon_2)}\cdot\Lambda^{\epsilon_1-1/2} \\
& \times p(\boldsymbol{\theta}_m,\boldsymbol{\theta}_p). \qquad (10)
\end{aligned}
$$

Spectral estimation by sampling from posterior distributions

If $0 < k \leqslant k_{max}$, by multiplying (7) and (9), the posterior distribution can be obtained as:

$$\begin{aligned} p(k, \mathbf{\Phi}|\mathbf{y}, 0 < k \leqslant k_{max}) \propto & (2\pi\sigma^2)^{-L/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}[\mathbf{y} - \mathbf{HDa}_k]^T[\mathbf{y} - \mathbf{HDa}_k]\right\} \\ & \times \frac{\Lambda^{k+\epsilon_1-1/2}}{k!} \exp[-\Lambda(1+\epsilon_2)] \cdot \frac{1}{|2\pi\sigma^2\mathbf{\Sigma}_k|^{1/2}} \cdot \exp\left[-\frac{\mathbf{a}_k^T\mathbf{\Sigma}_k^{-1}\mathbf{a}_k}{2\sigma^2}\right] \\ & \times \frac{\mathbb{I}_\Omega(k,\boldsymbol{\omega})}{\pi^k} \cdot \frac{\exp[-\gamma_0/(2\sigma^2)]}{\sigma^{2(v_0/2+1)}} \cdot (\delta^2)^{-\alpha_{\delta^2}-1} \cdot \exp\left(-\frac{\beta_{\delta^2}}{\delta^2}\right) \\ & \times p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p). \end{aligned} \tag{11}$$

After the integration of $\mathbf{a}_k$ and σ^2, this posterior distribution can be simplified as:

$$\begin{aligned} p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y}, 0 < k \leqslant k_{max}) \propto & (\gamma_0 + \mathbf{y}^T\boldsymbol{Q}_k\mathbf{y})^{-(1+L+v_0)/2} \times \frac{(\Lambda/((\delta^2+1)\pi))^k}{k!} \\ & \times \Lambda^{\epsilon_1-1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \cdot \delta^{-2(\alpha_{\delta^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta^2}}{\delta^2}\right] \\ & \times p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p), \end{aligned} \tag{12}$$

where

$$\boldsymbol{Q}_k = \boldsymbol{I}_{L\times L} - \mathbf{HD}\boldsymbol{R}\mathbf{D}^T\mathbf{H}^T \tag{13}$$

$$\boldsymbol{R}^{-1} = (1+\delta^{-2})\mathbf{D}^T\mathbf{H}^T\mathbf{HD} \tag{14}$$

$$\boldsymbol{r}_k = \boldsymbol{R}\mathbf{D}^T\mathbf{H}^T\mathbf{y}. \tag{15}$$

and $\boldsymbol{Q}_0 \triangleq \boldsymbol{I}_{L\times L}$, in which L is the length of the signal after being convolved with the receiver $\mathbf{H}$. In the case of $k = 0$, $\boldsymbol{a}_0^T\Sigma_0^{-1}\boldsymbol{a}_0 \triangleq 0$ and $|2\pi\sigma^2\Sigma_0|^{1/2} \triangleq 1$ are adopted in computation of the posterior distribution conventionally, as used in Andrieu and Doucet (1999); Copsey et al (2002). Therefore, the posterior distribution is presented as:

$$\begin{aligned} p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y}, k=0) \propto & \frac{1}{\sigma^{1+L}} \cdot \exp\left(-\frac{\mathbf{y}^T\mathbf{y} + \gamma_0}{2\sigma^2}\right) \\ & \times \frac{\Lambda^k/k!}{\pi^k} \cdot \Lambda^{\epsilon_1-1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \\ & \times \sigma^{-2(v_0/2+1)} \times p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p). \end{aligned} \tag{16}$$

In order to sample from this complicated posterior distribution, which takes account of the receiver characteristics, the procedure using a rjMCMC algorithm can be carried out. The algorithm splits the sampling of parameters into two main parts: a dimension-changing part and a fixed-dimension part. In general, in the former part, it samples the number of frequencies k from $p(k, \boldsymbol{\omega}_k|\mathbf{y}, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p)$; in the latter part, it samples the frequency values $\boldsymbol{\omega}_k$ from $p(\boldsymbol{\omega}_k|\mathbf{y}, k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p)$ and other model parameters in $\mathbf{\Phi}$. More specifically, in the process of dimension changing, there are two commonly used steps: a

birth move and a *death* move. The probability of choosing them are denoted as $b_{k^{(i)}}$ and $d_{k^{(i)}}$ respectively in each iteration i, which have already been defined in (Green, 1995). Furthermore, in the fixed dimension part, the algorithm includes an *update* move for a frequency with a probability $u_{k^{(i)}}$, an *update* move for the receiver parameters ($\boldsymbol{\theta}_m$, $\boldsymbol{\theta}_p$), and also an *update* move for other parameters, e.g. Λ and δ^2.

The acceptance ratio of the dimension-changing move is written as:

$$r_{change} = \frac{p(k+1, \boldsymbol{\Phi}_{k+1}|\mathbf{y})}{p(k, \boldsymbol{\Phi}_k|\mathbf{y})} \times \frac{q(k, \boldsymbol{\Phi}_k|k+1, \boldsymbol{\Phi}_{k+1})}{q(k+1, \boldsymbol{\Phi}_{k+1})|k, \boldsymbol{\Phi}_k)} \times \text{Jacobian}, \tag{17}$$

where $p(\cdot|\mathbf{y})$ is the posterior distribution of model parameters and $q(\cdot)$ is the proposal distribution for the parameters. Since either a *birth* move or a *death* move is a one step move, the Jacobian term will be one and thus has been omitted in (17).

In a *birth* move, the state of the constructed Markov chain changes from (k, $\boldsymbol{\Phi}_k$) at the iteration of i, to ($k+1$, $\boldsymbol{\Phi}_{k+1}$) at the iteration of $i+1$. A radial frequency ω^* is proposed uniformly on (0, π) and the new frequency vector becomes: $\boldsymbol{\omega}_{k+1} = [\boldsymbol{\omega}_k, \omega^*]$. In a *death* move, the state changes from ($k+1$, $\boldsymbol{\Phi}_{k+1}$) to (k, $\boldsymbol{\Phi}_k$). A frequency is randomly chosen to be removed from the existing $k+1$ frequency components. As a result, the proposal distributions for a *birth* move and a *death* move can be expressed as:

$$q(k+1, \boldsymbol{\Phi}^*_{k+1}|k, \boldsymbol{\Phi}_k) = b_k \times \frac{1}{\pi} \tag{18}$$

$$q(k, \boldsymbol{\Phi}_k|k+1, \boldsymbol{\Phi}^*_{k+1}) = d_{k+1} \div \begin{pmatrix} k+1 \\ 1 \end{pmatrix}. \tag{19}$$

Moreover, the ratio of proposal distributions in (18) and (19) can be simplified as:

$$\frac{q(k, \boldsymbol{\Phi}_k|\boldsymbol{k+1}, \boldsymbol{\Phi}_{k+1})}{q(k+1, \boldsymbol{\Phi}_{k+1}|k, \boldsymbol{\Phi}_k)} = \frac{\pi}{\Lambda}, \tag{20}$$

for $k = 0, \cdots, k_{max} - 1$. If $k = k_{max}$, the proposal ratio is zero.

Since $p(k, \boldsymbol{\omega}_k|\Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y}) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y})$, and using (12), the expression of the posterior ratio can be obtained as:

$$\frac{p(k+1, \boldsymbol{\omega}_{k+1}|\Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y})}{p(k, \boldsymbol{\omega}_k|\Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y})} = \left(\frac{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y}}{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_{k+1} \mathbf{y}}\right)^{1+L+v_0} \times \frac{\Lambda}{(k+1)(\delta^2+1)\pi}. \tag{21}$$

Accordingly, the acceptance ratio of the dimension-changing move can be achieved by substituting (20) and (21) into (17):

$$\begin{aligned} r_{change} &= \frac{p(k+1, \boldsymbol{\Phi}_{k+1}|\mathbf{y})}{p(k, \boldsymbol{\Phi}_k|\mathbf{y})} \times \frac{\pi}{\Lambda} \\ &= \left(\frac{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y}}{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_{k+1} \mathbf{y}}\right)^{1+L+v_0} \times \frac{1}{(k+1)(\delta^2+1)}. \end{aligned} \tag{22}$$

Furthermore, the probability of accepting a *birth* move and a *death* move can be presented as follows:

$$\alpha_{birth} = \min\{1, r_{change}\} \tag{23}$$
$$\alpha_{death} = \min\{1, r_{change}^{-1}\}. \tag{24}$$

In an *update* move for a frequency, the dimension of parameters are fixed. With a probability u_{k_i} at the ith iteration, the frequency values $\boldsymbol{\omega}_k$ are updated either based on the observations $\mathbf{y}$, or based on a random walk perturbation, using a hybrid Metropolis-Hasting (MH) step. The acceptance ratio of this move is expressed as:

$$r_{update} = \frac{p(\omega_{j,k}^*|k, \boldsymbol{\omega}_{-j,k}, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y})}{p(\omega_{j,k}|k, \boldsymbol{\omega}_{-j,k}, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y})} \times \frac{q(\boldsymbol{\omega}^*|\boldsymbol{\omega})}{q(\boldsymbol{\omega}|\boldsymbol{\omega}^*)}. \tag{25}$$

Since the full conditional posterior density can be stated as:

$$p(\omega_{j,k}|k, \boldsymbol{\omega}_{-j,k}, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p, \mathbf{y}) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y}), \tag{26}$$

and according to (12), the joint posterior density can be expressed as:

$$p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y}) \propto (\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y})^{-(1+L+v_0)/2}, \tag{27}$$

the acceptance ratio can be expanded as below:

$$r_{update} = \left(\frac{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y}}{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k^* \mathbf{y}}\right)^{(1+L+v_0)/2} \times \frac{q(\omega_{j,k}|\omega_{j,k}^*)}{q(\omega_{j,k}^*|\omega_{j,k})}. \tag{28}$$

In (28), $q(\cdot)$ denotes a proposal distribution for a single frequency. It is suggested that a hybrid MH step is used: $\lambda q_1(\cdot) + (1-\lambda) q_2(\cdot)$, in which $q_1(\cdot)$ represents a global update based on the frequency spectrum of the observations, and $q_2(\cdot)$ represents a local update using a random walk algorithm. Moreover, with the acceptance ratio of the *update* move, its probability of accepting this move is defined as:

$$\alpha_{update} = \min\{1, r_{update}\}. \tag{29}$$

If $k = 0$, the receiver parameters are sampled from their prior distribution $p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p)$. The demanding posterior distribution for updating the receiver parameters can be expressed as if $0 < k \leqslant k_{max}$:

$$p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p|k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y}) \propto p(k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \boldsymbol{\theta}_m, \boldsymbol{\theta}_p|\mathbf{y}). \tag{30}$$

Via (12), the distribution from which the samples are drawn can be presented as:

$$\begin{aligned} p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p|k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y}) \propto{}& (\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y})^{-(1+L+v_0)/2} \cdot \frac{\Lambda^k}{k! \pi^k} \cdot \left(\frac{1}{1+\delta^2}\right)^k \\ & \times \Lambda^{\epsilon_1 - 1/2} \cdot \exp[-\Lambda(1+\epsilon_2)] \times \delta^{-2(\alpha_{\delta^2}+1)} \cdot \exp\left[-\frac{\beta_{\delta^2}}{\delta^2}\right] \\ & \times p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p), \end{aligned} \tag{31}$$

where $\boldsymbol{Q}_k$ is the matrix stated in (13) and evaluated at the parameter values (k, $\boldsymbol{\omega}_k$, $\boldsymbol{\theta}_m$, $\boldsymbol{\theta}_p$, δ^2).

The priors for $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$ are both assigned Gaussian distributions centered about these rough estimates. Moreover, they are assumed to be independent with each other: $p(\boldsymbol{\theta}_m, \boldsymbol{\theta}_p) = p(\boldsymbol{\theta}_m) \cdot p(\boldsymbol{\theta}_p)$. Single terms in $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$ are also independent with one another, such that,

$$p(\boldsymbol{\theta}_m) = p(\theta_{m_1}) p(\theta_{m_2}) \cdots p(\theta_{m_P}) \tag{32}$$
$$p(\boldsymbol{\theta}_p) = p(\theta_{p_1}) p(\theta_{p_2}) \cdots p(\theta_{p_P}). \tag{33}$$

Therefore, the prior for each single term in $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$ can be expressed as:

$$\theta_{m_i} \sim \mathcal{N}(\theta_{m_i}^{(init)}, \sigma_m^2) \tag{34}$$
$$\theta_{p_i} \sim \mathcal{N}(\theta_{p_i}^{(init)}, \sigma_p^2), \tag{35}$$

where i denotes the index in vector $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$. σ_m^2 and σ_p^2 are assigned constants.

In order to sample from the distributions $p(\boldsymbol{\theta}_m | k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y})$ and $p(\boldsymbol{\theta}_p | k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y})$, standard MH steps are used. As θ_{m_i} and θ_{p_i} are the ith item in the vector $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_p$, and they are independent with one another, the following formulas are suitable for updating each single item. The acceptance probability for updating one item, θ_i, in $\boldsymbol{\theta}_m$ or $\boldsymbol{\theta}_p$, can be stated as:

$$\alpha_\theta(\theta_i, \theta_i^*) = \min\{1, r_\theta(\theta_i, \theta_i^*)\}, \tag{36}$$

where $r_\theta(\theta_i, \theta_i^*)$ represents the acceptance ratio for updating one of the receiver parameters θ_i, obtained from (31):

$$\begin{aligned} r_\theta(\theta_i, \theta_i^*) &= \frac{p(\theta_i^* | k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y})}{p(\theta_i | k, \boldsymbol{\omega}_k, \Lambda, \delta^2, \mathbf{y})} \times \frac{q_\theta(\theta_i | \theta_i^*)}{q_\theta(\theta_i^* | \theta_i)} \\ &= \left(\frac{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y}}{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k^* \mathbf{y}} \right)^{(1+L+v_0)/2} \times \frac{p(\theta_i^*)}{p(\theta_i)} \times \frac{q_\theta(\theta_i | \theta_i^*)}{q_\theta(\theta_i^* | \theta_i)}, \end{aligned} \tag{37}$$

where $\boldsymbol{Q}_k$ and $\boldsymbol{Q}_k^*$ are matrices stated in (13) for receiver parameters ($\boldsymbol{\theta}_m$, $\boldsymbol{\theta}_p$) and ($\boldsymbol{\theta}_m^*$, $\boldsymbol{\theta}_p^*$), respectively.

In (37), $q_\theta(\cdot)$ is the proposal distribution for one item in the receiver parameters, θ_i. For ease of the computation, it is chosen as a random walk on the space of the polynomial coefficients in the magnitude response or the phase response. In other words, each proposed coefficient depends on its previous value, the update is:

$$\theta_i^* | \theta_i \sim \mathcal{N}(\theta_i, \sigma_\theta^2), \tag{38}$$

where σ_θ^2 is a constant that determines the step size in updating θ_i.

The rest of the parameters in the signal model can be categorized as hyperparameters, Λ and δ^2, and nuisance parameters $\mathbf{a}_k$ and σ^2. Generally speaking, there is no need to update the nuisance parameters $\mathbf{a}_k$ and σ^2 as they have already been integrated out from the joint posterior distribution. Unfortunately, they are required when updating

Table 3
Update the parameters

Parameters	Prior Distribution	Posterior Distribution
$\mathbf{a}_k$	$\mathcal{N}(0, \delta^2 \mathbf{D}^T \mathbf{H}^T \mathbf{H} \mathbf{D})$	$\mathcal{N}(\boldsymbol{r}_k, \sigma^2 \boldsymbol{R})$
σ^2	$\mathcal{IG}(v_0/2, \gamma_0/2)$	$\mathcal{IG}(\frac{1+L+v_0}{2}, \frac{\gamma_0 + \mathbf{y}^T \boldsymbol{Q}_k \mathbf{y}}{2})$
Λ	$\mathcal{G}a(1/2 + \epsilon_1, \epsilon_2)$	$\mathcal{G}a(1/2 + k + \epsilon_1, 1 + \epsilon_2)$
δ^2	$\mathcal{IG}(\alpha_{\delta^2}, \beta_{\delta^2})$	$\mathcal{IG}(k + \alpha_{\delta^2}, \frac{1}{2\sigma^2 (\mathbf{HDa}_k)^T (\mathbf{HDa}_k)} + \beta_{\delta^2})$

Table 4
Parameter settings for the simulated signal.

k	$E_k = a_{s,k}^2 + a_{c,k}^2$	$-\arg\tan(a_{s,k}/a_{c,k})$	ω_k
1	20	0	0.2π
2	6.3246	$\pi/4$	0.4π
3	20	$\pi/3$	0.6π

hyperparameters Λ and δ^2. Therefore, all of them require being updated in every single iteration. Furthermore, the priors for them are chosen as conjugate priors so as to make the computation much easier. Their priors and the corresponding posteriors are stated in Table 3.

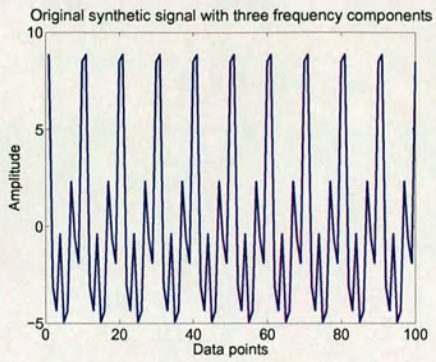
Results

Evaluation on a simulated signal

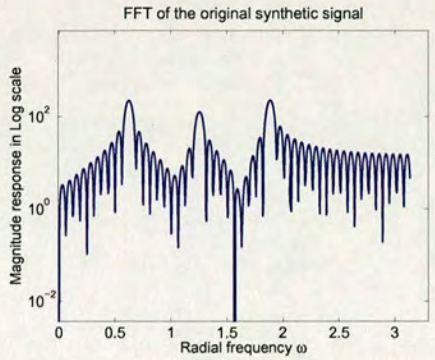
The simulated signal has a length of $N = 100$ data points, and has three frequency components: 0.2π, 0.4π and 0.6π. Its parameter setting is displayed in Table 4. The original simulated signal and its frequency spectrum are shown in Figure 6(a) and Figure 6(b) respectively. Then the original synthetic signal convolves with the impulse response of the receiver, which is shown in Figure 7. The obtained filtered synthetic signal after being convolved with the calibrated ultrasound receiver in the time domain is shown in Figure 6(c) and the signal-to-noise ratio (SNR) is 20dB. Its corresponding frequency spectrum is also shown in Figure 6(d).

Furthermore, as the calibrated receiver has a length of $M = 256$ data points, the filtered synthetic signal has a length of $L = N + M - 1 = 355$ data points. Since the magnitude and the phase responses are modeled as polynomial functions, the initial estimated values for their coefficients $\boldsymbol{\varphi}_m^{(init)}$ and $\boldsymbol{\varphi}_p^{(init)}$ are given in Table 5.

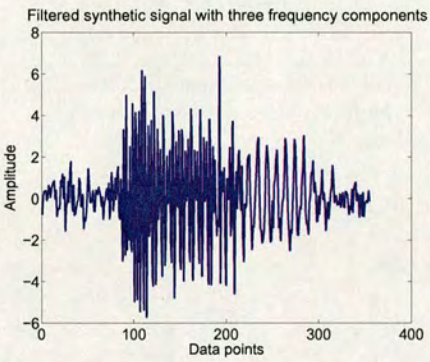
Consider the simulated signal with a single noise realisation, a Markov chain, is constructed. After a certain number of iterations, samples from the desired posterior distribution can be drawn. Then the MAP estimator is used to estimate the model parameters, especially the parameters of interests, i.e. the number of frequencies k and their corresponding values $\boldsymbol{\omega}_k$. Seen from Figure 8(a), the constructed Markov chain converged



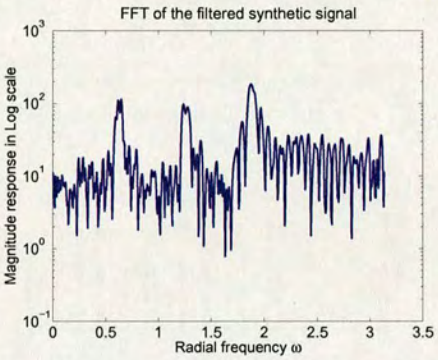
(a) The original simulated noisy free signal in the time domain.



(b) The original simulated noisy free signal in the frequency domain.



(c) Filtered simulated signal in the time domain.



(d) Filtered simulated signal in the frequency domain.

Figure 6. The original noise free signal and the filtered signal with the calibrated ultrasound receiver in both the time and the frequency domain.

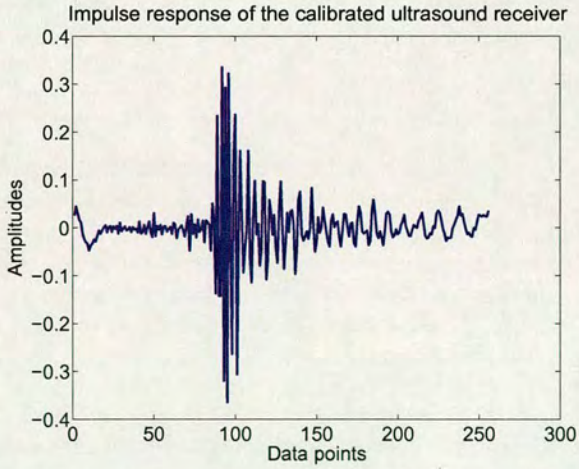


Figure 7. Impulse response of the calibrated ultrasound receiver.

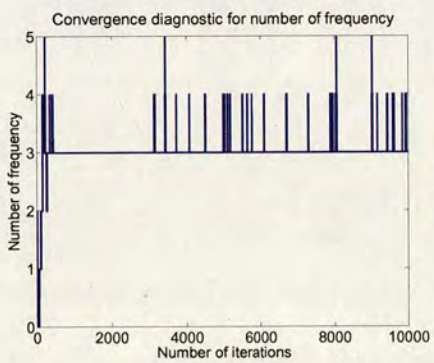
Table 5
Initial values for the coefficients of the magnitude and the phase responses of the receiver

$\varphi_m^{(init)}$	Coefficient values
$\hat{\varphi}_{m_1}$	-180.6035
$\hat{\varphi}_{m_2}$	505.7294
$\hat{\varphi}_{m_3}$	-578.6615
$\hat{\varphi}_{m_4}$	352.5759
$\hat{\varphi}_{m_5}$	-124.9146
$\hat{\varphi}_{m_6}$	26.5174
$\hat{\varphi}_{m_7}$	-3.3228
$\hat{\varphi}_{m_8}$	0.2266
$\hat{\varphi}_{m_9}$	-0.0065

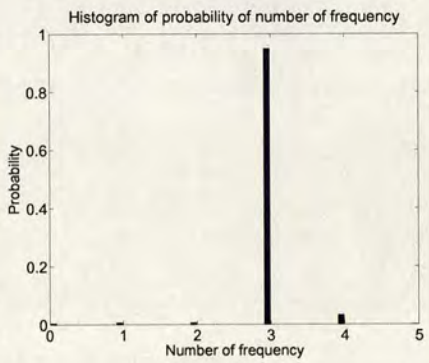
(a) Initial coefficient values for the magnitude response of the receiver

$\varphi_p^{(init)}$	Coefficient values
$\hat{\varphi}_{p_1}$	$-1.984e+3$
$\hat{\varphi}_{p_2}$	0.0078
$\hat{\varphi}_{p_3}$	$-1.2204e-8$
$\hat{\varphi}_{p_4}$	$9.9765e-15$
$\hat{\varphi}_{p_5}$	$-4.6898e-21$
$\hat{\varphi}_{p_6}$	$1.3206e-27$
$\hat{\varphi}_{p_7}$	$-2.1938e-34$
$\hat{\varphi}_{p_8}$	$1.9747e-41$
$\hat{\varphi}_{p_9}$	$-7.3994e-49$

(b) Initial coefficient values for the phase response of the receiver



(a) Convergence diagnostic for number of frequency.



(b) Histogram of number of frequency.

Figure 8. Estimation of number of frequency based on the signal model with one noise realisation.

quickly and the burn-in period is around 1000 iterations. By discarding the first 1000 samples from its equilibrium distribution, the rest of 9000 samples are used to plot a histogram of the number of frequency in the original signal. Figure 8(b) displays this histogram, indicating the estimated number of frequency in the original unfiltered signal is three, with the highest probability. Moreover, Figure 9 displays histograms of the estimated frequency values. The peak values in the histograms indicate the most probable frequency values in the original signal.

Furthermore, there are 100 different noise realisations used to check the stability of the developed rjMCMC algorithm. Table 6 shows the spread of MAP estimates of the number of frequency for the original signal using 100 Monte Carlo runs. It can be inferred that 91% of the time, the detected number of frequency is 3. Moreover, the estimated frequency values are compared to the ground truths, which are shown in Table 7. The

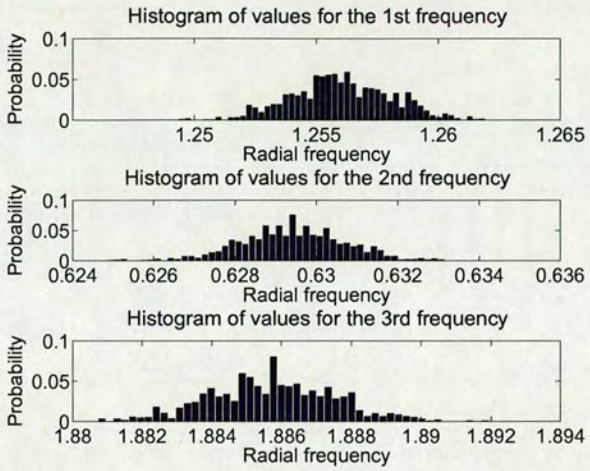


Figure 9. Histograms of frequency values based on the signal model with one noise realisation.

Table 6
Spread of MAP estimates of number of frequency for the original signal with 100 noise realisations

K	$0 \sim 2$	3	4	5	$\geqslant 6$
MAP	0	91	8	1	0

Table 7
Comparison of frequency values between the ground truth and the estimates

Frequency components	Ground Truth	Estimated values (mean $\pm$ SD)
1^{st} Freq.	$0.2\pi = 0.6283$	0.6323 ± 0.0099
2^{nd} Freq.	$0.4\pi = 1.2566$	1.2451 ± 0.0877
3^{rd} Freq.	$0.6\pi = 1.8850$	1.8723 ± 0.0895

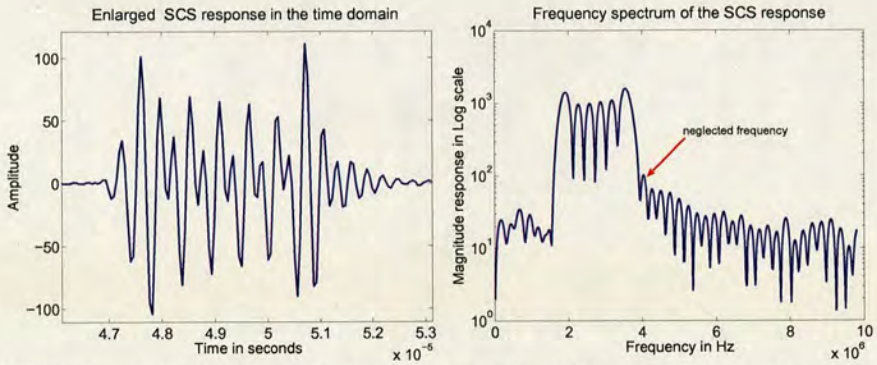
estimated means and SD values are also presented. The errors between the means and the ground truths are small, and the small SD values indicate good stability of the developed estimation algorithm.

Estimation of measured ultrasound echo signals

With this proposed model and its corresponding rjMCMC estimation algorithm, the true spectral contents of measured ultrasound echo signals can be recovered and extracted correctly. Therefore, the real behavior of ultrasound scatterers will be revealed.

Estimation of an SCS response

In the frequency spectrum of measured ultrasound echo signals from SCSs, due to the bandwidth limit of the ultrasound receiver, the frequencies outside the frequency range, i.e. from 1MHz to 5MHz, will be attenuated and thus might be regarded as side lobes



(a) Enlarged SCS response in the time domain.

(b) FFT of SCS response in the frequency domain.

Figure 10. Display of the measured SCS response in the time and the frequency domains.

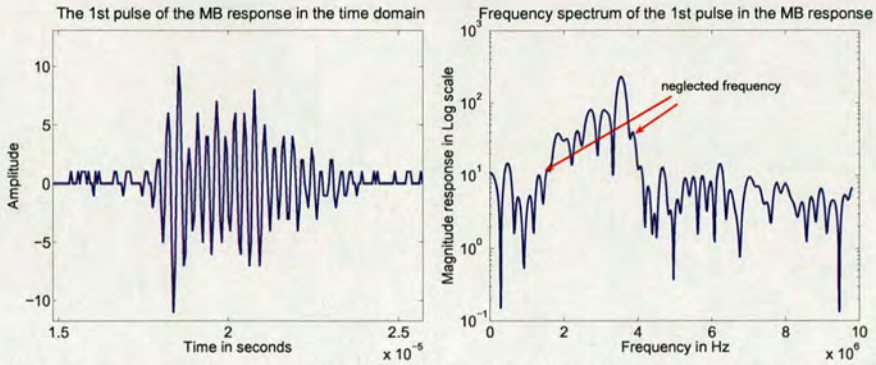
Table 8
Spread of MAP estimates of number of frequency in the SCS response with 100 repeats of the algorithm.

K	$0 \sim 1$	2	3	4	5	$\geqslant 6$
MAP	0	4	10	40	46	0

Table 9
Estimated frequency values of the measured SCS response

Freq. components	Means (in MHz)	SD values (in MHz)
1st Freq.	1.6772	0.3876
2nd Freq.	2.3684	0.5427
3rd Freq.	3.0139	0.5653
4th Freq.	3.4766	0.1926
5th Freq.	3.8425	0.7956

rather than the main lobe peaks. Take one SCS response with a transmit frequency, which has a fundamental frequency at 1.83MHz as an example. Figure 10(a) displays an enlarged version of the SCS response in the time domain. The developed spectral estimation algorithm is repeatedly applied to this SCS response for 100 times. The spread of the MAP estimates of the number of frequency is given in Table 8. Moreover, the estimated frequency values are displayed in Table 9. Their means and SD values can be computed based on the results from 100 replications of the developed rjMCMC estimation algorithm. Compared to the frequency spectrum of this SCS response in Figure 10(b), one frequency component detected by the estimation algorithm, 3.8425MHz, has a very small amplitude in the spectrum. Due to the attenuation introduced by the ultrasound receiver, it was regarded as a sidelobe instead of a mainlobe peak in the spectrum. However, it has been detected when the proposed model is assumed and the developed estimation algorithm is used. Moreover, the obvious four peaks between 2.0MHz and 4.0MHz in the spectrum are suppressed when the receiver bandwidth is considered. Only two of them are regarded as main lobes and the other two peaks are neglected as side lobe noise using the developed rjMCMC estimation algorithm.



(a) Enlarged version of the 1^{st} pulse of the MB response in the time domain.

(b) FFT of the 1^{st} pulse of the MB response in the frequency domain.

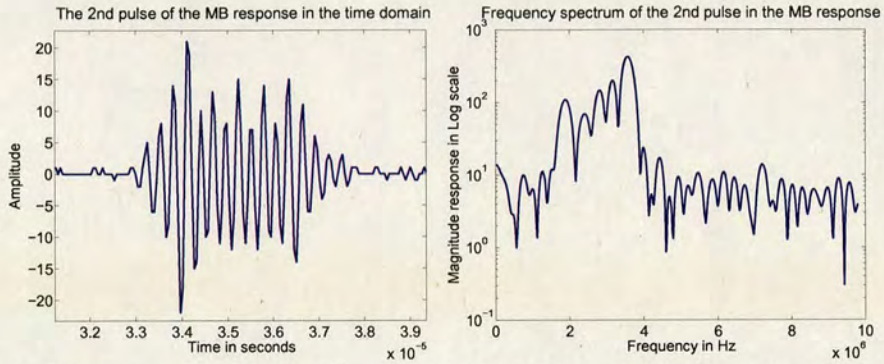
Figure 11. Display of the 1^{st} pulse of the measured MB response in the time and the frequency domains.

Estimation of an MB response

With respect to measured ultrasound echo signals from MBs, if the receiver characteristics are not taken into consideration, their spectra will be in some distortion. This distortion may lead to a misunderstanding of MB behavior due to the incorrect spectral estimation. Therefore, the rjMCMC estimation algorithm based on the proposed signal model is trying to recover the original spectra of MB responses from the distortion and finally reveal the true MB behavior from experimental measurements. Take one MB response with an incident pulse, which has a fundamental frequency at 1.83MHz as an example. There are two pulse segments in the MB response. They are displayed in Figure 11(a) and Figure 12(a) respectively. As the proposed signal model and the developed estimation algorithm are only valid for a single pulse duration, these two pulse segments in the MB response are separated manually and then are estimated individually. The same as the estimation performed for the SCS response, the algorithm is carried out 100 times on two pulses in the MB response. Table 10 displays the spread of the MAP estimates of the number of frequency for two MB pulse segments. The results indicate there are 7 frequency components in the first pulse and 5 in the second pulse as they have the highest probabilities. Furthermore, the means of the estimated frequency values for two pulses are presented in Table 11, as well as their corresponding SD values. These two MB pulse segments have similar spectral contents although their signal shape in the time domain are different. Moreover, they have more harmonics than those in the SCS response, as predicted in ultrasonics literature.

Discussion

This paper has proposed a signal model that takes the receiver characteristics as model parameters, and has developed a rjMCMC estimation algorithm to extract the spectral contents in measured ultrasound echo signals. The proposed model and the developed estimation algorithm are shown to be robust in estimating the spectral content of ultrasound scatterers. Some frequency components, which are suppressed by the limited



(a) Enlarged version of the 2^{nd} pulse of the MB response in the time domain.

(b) FFT of the 2^{nd} pulse of the MB response in the frequency domain.

Figure 12. Display of the 2^{nd} pulse of the measured MB response in the time and the frequency domains.

Table 10
Spread of MAP estimates of number of frequency for two pulse segments in the MB response

K_1	$0 \sim 3$	4	5	6	7	8	$\geqslant 9$
MAP	0	2	8	28	60	2	0

(a) Spread of MAP estimates of number of frequency for the 1^{st} pulse

K_2	$0 \sim 1$	2	3	4	5	6	$\geqslant 7$
MAP	0	2	2	32	38	26	0

(b) Spread of MAP estimates of number of frequency for the 2^{nd} pulse

bandwidth of the ultrasound receiver, can be recovered. Although this paper only considers a single pulse segment in the ultrasound responses, it is easy for the signal model and the estimation algorithm to extend to include all pulse segments in multiple pulse echo signals.

From the perspective of ultrasound contrast imaging, according to the analysis of the spectral content of measured ultrasound echo signals, the real behavior of scatterers, especially the MBs, have been revealed . This is able to broaden the research area of ultrasound MBs. Furthermore, it will also assist in discrimination of echo signals from MBs and SCSs.

Acknowledgment

The authors would like to thank Robin Steel for providing codes of theoretical responses from SCSs. This work is partly funded by Biologically inspired acoustic systems (BIAS).

Table 11
Estimated frequency values for two pulse segments in the MB response

Freq. components	Means (in MHz)	SD values (in MHz)
1^{st}	1.6869	0.5187
2^{nd}	2.0637	0.3797
3^{rd}	2.5152	0.4463
4^{th}	2.9093	0.3228
5^{th}	3.1695	0.2719
6^{th}	3.4006	0.2398
7^{th}	3.8942	0.2456

(a) Estimated frequency values for the 1^{st} pulse

Freq. component	Means (in MHz)	SD values (in MHz)
1^{st}	1.8478	0.0623
2^{nd}	2.5615	0.3464
3^{rd}	2.9011	0.2867
4^{th}	3.3241	0.2791
5^{th}	3.6147	0.1601

(b) Estimated frequency values for the 2^{nd} pulse

Figure captions

Figure 1 The procedure of extraction of the real MB behavior.
Figure 2 A simple illustration of the receiver calibration.
Figure 3 Receiver calibration in the whole frequency range.
Figure 4 Curve fitting for individual experimental data with known model order.
Figure 5 Averaged phase response obtained based on the RTF with 21 different transmit frequencies.
Figure 6 The original noise free signal and the filtered signal with the calibrated ultrasound receiver in both the time and the frequency domain.
Figure 7 Impulse response of the calibrated ultrasound receiver.
Figure 8 Estimation of number of frequency based on the signal model with one noise realisation.
Figure 9 Histograms of frequency values based on the signal model with one noise realisation.
Figure 10 Display of the measured SCS response in the time and the frequency domains.
Figure 11 Display of the 1^{st} pulse of the measured MB response in the time and the frequency domains.
Figure 12 Display of the 2^{nd} pulse of the measured MB response in the time and the frequency domains.

References

Andrieu C, Doucet A. Joint bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. 1999;47:2667–2676.

Copsey K, Gordon N, Marrs A. Bayesian analysis of generalized frequency-modulated signals. 2002;50:725–735.

de Jong N, Hoff L, Skotland T, Bom N. Absorption and scatter of the encapsulated gas filled microspheres: theoretical considerations and some measurements. Ultrasonics. 1992;30:95–103.

Frinking PJA, de Jong N. Acoustic modelling of shell-encapsulated gas bubbles. Ultrasound Med.Biol.. 1998;24:523–533.

Green PJ. Reversible jump MCMC computation and bayesian model determination. Biometrika. 1995;82:711–732.

Haykin S, Adaptive filter theory. Prentice Hall. 2002.

Hilgenfeldt S, Lohse D, Zomack M. Sound scattering and localized heat deposition of pulse-driven microbubbles. Journal of the Acoustical Society of America. 2000;107(6):3530–3539.

Krishna PD, Shankar PM, Newhouse VL. Subharmonic generation from ultrasonic contrast agents. Phys Med Biol. 1999;44:681–694.

Manolakis DG, Ingle VK, Kogon SM, Statistical and adaptive signal processing: Spectral estimation, signal modeling, adaptive filtering and array processing. McGrawHill Companies, Inc.. 2000.

Marsh JN, Hughes MS, et al. CSH. Frequency and concentration dependence of the backscatter coefficient of the ultrasound contrast agent. Albunex.J Ac Soc Am. 1998;104:1654–1666.

Ruanaidh JJKO, Fitzgerald WJ, Numerical Bayesian methods applied to signal processing. Springer. 1996.

Sboros V, Pye SD, MacDonald CA, et al. Absolute measurement of ultrasonic backscatter from single microbubbles. Ultrasound in Med.&Biol.. 2005;31:1063–1072.

Digital Object Identifier: 10.1109/ULTSYM.2008.0335

Statistical spectral analysis for echo signals from microbubbles and solid spheres

Yan Yan, James R. Hopgood
Institute for Digital Communications
School of Engineering and Electronics
University of Edinburgh, Edinburgh, U.K.
{Y.Yan, James.Hopgood}@ed.ac.uk

Robin Steel, Vassilis Sboros
Department of Medical Physics
School of Clinical Sciences and Community Health
University of Edinburgh, Edinburgh, U.K.
{Robin.Steel, Vassilis.Sboros}@ed.ac.u

Abstract—**The spectral analysis of echo signals from linear and non-linear scatters is important for understanding the behavior of ultrasound contrast microbubbles (MBs) and assisting the design of insonifying pulse waveforms. Most pulse designs assume a theoretical MB behavior instead of using experimental data despite it being accepted that theoretical models have not succeeded in describing MB behavior. Fourier transform (FT) based non-parametric spectral analysis methods are widely used but have limitations in spectral and temporal localization. This study is the first to incorporate receiver characteristics and introduce a parametric model for the estimation of temporal and spectral content of experimental echo signals. This study improves spectral analysis performance and helps understand MB behavior.**

Keywords: ***ultrasound contrast microbubbles, spectral analysis, parametric estimation, receiver chracteristics***

I. INTRODUCTION

The spectral analysis of echo signals from microbubbles (MBs) used as Ultrasound Contrast Agents (UCAs) [1] can provide much information about MB behavior. The frequency content provides very crucial features for differentiating the responses from MBs and tissue. Within a number of research fields, including digital and statistical signal processing, spectral analysis has been a popular topic since the development of Fourier transform (FT) and specifically the fast Fourier transform (FFT). However, the Fourier transform does not incorporate any prior knowledge regarding the structure or content of the signals. The failure to incorporate such knowledge means there is a fundamental trade-off between temporal and spectral localization, as stated by the time-frequency uncertainty principle. Therefore, for short imaging pulses, there can be a significant overlap between spectral components when analyzed using the Fourier transform. Similarly, in MB detection algorithms dependent on harmonic structure, standard linear time-invariant (LTI) filters are unable to perfectly extract the nonlinear MB signal component from linear tissue signals. Harmonic analysis has moved on in the field of statistical signal processing with the development of Bayesian spectral analysis [2]. In the Bayesian framework, a parametric model is proposed which incorporates prior information regarding the signal structure. However, as with all parametric methods, the number of spectral components present in the signal needs to be estimated. This thus becomes a model order selection problem. There are several standard well known methods for model order selection, namely Akaike information criterion (AIC) [3] and the minimum description length (MDL) [4]. However, the Bayesian framework facilitates joint parameter estimation and model order selection. Andrieu and Doucet [5] proposed a joint Bayesian model selection algorithm for frequency estimation which outperforms the techniques based on AIC and MDL. Since the reflected signals from MB and tissue have low signal-to-noise ratio (SNR) and short pulse length, Bayesian model selection offers more reliable estimates of the spectral content and number of frequencies present in the signal.

Ultrasound transducers have a band-limited property. Thus, the parametric Bayesian frequency estimation technique proposed in [5] must be modified to incorporate the receiver characteristics. There has been many investigations into removing the influence of receiver characteristics [6] in order to obtain more accurate knowledge about the MB behavior. The simplest approach to removing the effect of receive filter characteristics is to design an inverse filter. In this paper, a robust temporal and spectral estimation algorithm is presented in which the receiver response is accounted. The method is described in Section II and the complete Bayesian algorithm is presented in Section III. In Section IV, the performance of the newly proposed estimation algorithm is evaluated by analyzing first synthetic data where the groundtruth is known, followed by real experimental data. Conclusions are drawn in Section V.

II. METHOD

A. Receiver Characteristics

MBs are non-linear scatters, and the theory available for describing MBs is not sufficient to predict the MB behavior at the receiver input. Thereby, real MB behavior can only be achieved by analyzing experimental responses. Thus, it is crucial to account for the way in which the receiver characteristics modify the actual MB at response. The receiver response can be estimated by looking at the signals at the input and output of the transducer (Fig 1.). While it is possible to measure the transmitted pulse of the ultrasound emitter, it is difficult to measure the signal at the input to the ultrasound transducer. The approach used in [6] therefore excites small linearly scattering solid copper spheres for which the returned signal can, unlike nonlinear MB signals, be predicted using a theoretical physical model *given* the transmitted waveform.

This work is sponsored by the EPSRC under their basic technology grant, "Biologically Inspired Acoustic Systems", EP/C523776/1.

978-1-4244-2480-1/08/$25.00 ©2008 IEEE

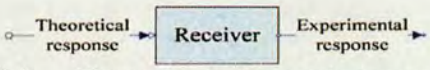


Fig.1. Diagram of receiver impact on the responses

The magnitude response of the receiver can be calibrated by taking the ratio of the energy density of the experimental and the theoretical responses from the solid spheres (see Fig. 1). Since the transmit pulse is a narrowband signal containing a dominant fundamental and 2nd-harmonic spectral component, the ratio of the magnitude responses will only cover a limited bandwidth in the spectrum. To estimate the full receiver spectral response, it is necessary to use multiple transmit pulses with different frequencies [6]; for each transmit frequency, the corresponding gain of magnitude response of the receiver is evaluated. Thus, the frequency response of the transducer is estimated by filtering the transmit pulse into fundamental and 2nd-harmonic signals, calculating the theoretical pulse echo return at the transducer input, filtering the measured transducer input in the same way, and then calculating the ratio of the energy density of the output to input. However, this procedure yields an estimated sampled spectrum, and does not provide a continuous receiver response curve that may allow the design of an inverse filter from individual calibration values.

B. *Inverse filter design*

The receiver magnitude response calculated by Sboros [6] is only available for a finite number of frequency values, f_k, depicted as circles in Fig.2. These samples can be interpolated to obtain a continuous response using a curve fitting technique.

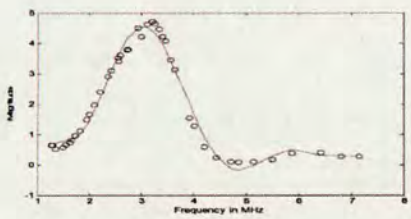


Fig.2. Receiver magnitude frequency response.

A common approach to fitting this curve is the least square (LS) method. The magnitude response, f_k, as a function of frequency bin, k, is modeled as a polynomial with unknown coefficients:

$$f_k = \sum_{p=1}^{P} c_p k^p ,$$

These coefficients can be evaluated by solving the standard least squares equations. However, to obtain uncertainty in the frequency response, it is necessary to use numerical methods by sampled from the posterior density, which is a multivariate normal distribution, by Gibbs sampler [7]. The resulting curve for the full magnitude frequency response is shown in Fig. 2.

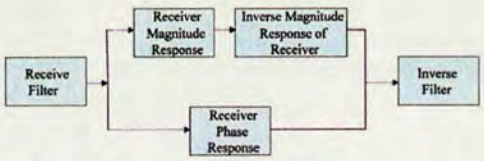


Fig.3. Block diagram of inverse filter design procedure

Fig.3 shows a block diagram of the inverse filter design procedure. Since the phase response cannot easily be obtained by experiments, it is unmodified. The resulting inverse magnitude response is shown in Fig.4.

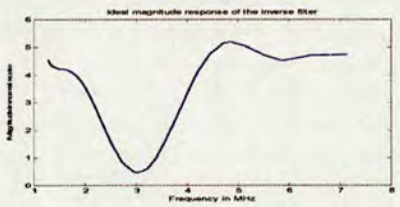


Fig.4. Inverse magnitude response of the receiver

The signal measured at the output of the receiver transducer is now modified by de-convolving with the inverse filter to obtain an estimate of the original FFT spectrum. This removes the influence of receiver characteristics. This modified signal is now analyzed using a modified version of the Bayesian spectral analysis algorithm proposed in [5]. Accordingly, it will be seen more accurate MB behavior can be obtained on the basis of correct estimation of frequency components.

III. ESTIMATION ALGORITHM

The spectral estimation algorithm proposed here is a Bayesian parametric estimation method. It requires a signal model and priors for all parameters. The estimation of the desired parameters is achieved by finding the maximum of the a-posterior distribution: the so-called MAP estimate. However, in this case, the posterior distribution for the frequency content is non-linear and there is no closed-form solution for the MAP estimate. Therefore, Markov chain Monte Carlo is used, and a reversible jump technique is used to select the order of signal model.

A. *Signal Model*

The responses from nonlinear MB scatterers and linear scattering solid spheres data can both be modeled as the sum of finite duration sinusoids. The vector-matrix form is:

$$\mathbf{y} = \mathbf{D}(\boldsymbol{\omega}_k)\mathbf{a}_k + \mathbf{n} . \qquad (1)$$

where $\mathbf{y}$ is the filtered response vector of N observation samples of the filtered pulse echo return, and $\mathbf{n}$ is a zero-mean

white Gaussian noise (WGN) vector representing a modeling error. In addition, $\mathbf{a}_k$ is the amplitude vector, and $\mathbf{D}(\boldsymbol{\omega}_k)$ is a matrix with frequency and pulse location information:

$$\mathbf{D}(\boldsymbol{\omega}_k)=\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ E(\omega_1 t(T_1)) & \ddots & E(\omega_k t(T_1)) \\ \vdots & \cdots & \vdots \\ E(\omega_1 t(T_2-1)) & \ddots & E(\omega_k t(T_2-1)) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

Here, $E(\cdot)=[\cos(\cdot),\sin(\cdot)]$. Moreover, T_1 and T_2 represent the beginning and end points of the pulse, $\omega_1,\cdots,\omega_k$ denote k different frequency components. The number of sinusoids k and their parameters $\boldsymbol{\theta}_k \triangleq (T_1,T_2,\boldsymbol{\omega}_k,\mathbf{a}_k,\sigma^2)$ are all unknown. The objective is to estimate k, T_1,T_2 and $\boldsymbol{\omega}_k$. The likelihood function can be obtained based on the proposed signal model and prior information, as expressed in (2):

$$p(\mathbf{y}\,|\,k,\boldsymbol{\theta}_k)=(2\pi\sigma^2)^{-N/2} \times\exp\left\{-\frac{1}{2\sigma^2}[\mathbf{y}-\mathbf{D}(\boldsymbol{\omega}_k)\mathbf{a}_k]^T\times[\mathbf{y}-\mathbf{D}(\boldsymbol{\omega}_k)\mathbf{a}_k]\right\}. \quad (2)$$

B. Priors and Posteriors of Unknown Parameters

The prior distribution of the unknown parameters reflects the degree of belief of the relevant values of the parameters.

$$p(k,\boldsymbol{\theta}_k)=p(k,\mathbf{a}_k,\boldsymbol{\omega}_k)p(\sigma^2)\propto\frac{\Lambda^k}{k!}\cdot\exp(-\Lambda) \times\frac{1}{|2\pi\sigma^2\Sigma_k|^{1/2}}\cdot\exp\left[-\frac{\mathbf{a}_k^T\Sigma_k^{-1}\mathbf{a}_k}{2\sigma^2}\right]\cdot\frac{1}{\pi^k}\cdot\frac{1}{\sigma^2} \quad (3)$$

where $\Sigma_k^{-1}=(1+\delta^{-2})\mathbf{a}_k^T\mathbf{a}_k$. Bayes's theorem is used to obtain the joint posterior distribution after the nuisance parameters $\mathbf{a}_k$ and σ^2 are integrated out [5, 7], to give:

$$p(k,\boldsymbol{\omega}_k,T_1,T_2\,|\,\mathbf{y})\propto(\gamma_0+\mathbf{y}^T\mathbf{P}_k\mathbf{y})^{-(N+\nu_0)/2}\cdot\frac{(\Lambda/[(\delta^2+1)\pi])^k}{k!}. \quad (4)$$

where $\mathbf{P}_k=\mathbf{I}_N-\mathbf{D}(\boldsymbol{\omega}_k)\mathbf{M}_k\mathbf{D}^T(\boldsymbol{\omega}_k)$, $\mathbf{I}_N$ is a $N\times N$ identity matrix and $\mathbf{M}_k^{-1}=\mathbf{D}^T(\boldsymbol{\omega}_k)\mathbf{D}(\boldsymbol{\omega}_k)+\Sigma_k^{-1}$. $\nu_0,\gamma_0,\Lambda,\delta^2$ are hyper-parameters of the Bayesian model. Note that the posterior in equation (4) is highly non-linear and thus reversible jump Markov chain Monte Carlo method is adopted to facilitate sampling from the posterior distribution [8].

C. Reversible jump Markov chain Monte Carlo (rjMCMC)

In the standard MCMC algorithm, an ergodic Markov chain for the unknown parameters $(k^{(i)},\boldsymbol{\theta}_{k^{(i)}})$ is constructed whose equilibrium distribution is the specified joint posterior distribution in (4). If the simulation is run sufficiently long, the drawn samples reach a stationary distribution equivalent to the posterior. These samples are used to provide an empirical approximation to the posterior density in (4) and provide, for example, a MAP estimate. Combining a reversible jump technique, the reversible jump MCMC (rjMCMC) algorithm can also provide a solution to the automatic model order selection problem. A MAP estimator can also be used to determine the value of k, i.e. the number of frequency component, which maximizes $p(k\,|\,\mathbf{y})$. There are three main components of the rjMCMC algorithm for model selection, as shown in Fig.5 which gives a simple illustration for a signal with 3 frequency components.

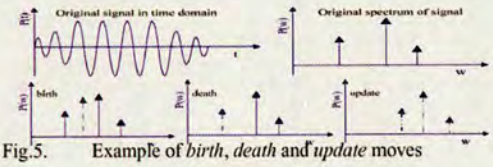


Fig.5. Example of *birth*, *death* and *update* moves

- *Birth* move: with probability b_k proposes a new frequency component in the normalized frequency range $[0,\pi]$ and changes the number of frequency components in the signal model from k to $k+1$.
- *Death* Move: with probability d_k removes an existing frequency component randomly and changes the number of frequency components from k to $k-1$.
- *Update* Move: with probability u_k updates all the amplitude and frequency parameters for all k frequencies. The number of components remains unchanged.

In the update move, the frequency parameters are either updated based on the observations, or as a perturbation of the existing parameter values, using a hybrid Metropolis-Hasting step [8] where the invariant distribution function is given by, $p(\boldsymbol{\omega}_{j,k}\,|\,\mathbf{y},\boldsymbol{\omega}_{-j,k})$, $(\boldsymbol{\omega}_{-j,k}\triangleq[\omega_{1,k},\cdots,\omega_{j-1,k},\omega_{j+1,k},\cdots])$. The update is thus performed using a proposal density as follows:

- *Update* based on observations with probability λ. The proposal distribution is the modified FFT spectrum discussed in Section II, i.e., taking account the receiver characteristics and the inverse filter design. The basic idea is to propose a new frequency, independent of previous one, and explore the regions around obvious peaks in the modified spectrum.
- *Update* by random perturbation: with probability $1-\lambda$. The proposal distribution is based on a random walk.

In the time domain, the estimation of the pulse locations is of great importance for analyzing the reflected signals. A ssingle pulse echo signals are considered here, the number of

change-points (related to beginning and endpoints of the pulse segments, also known as the "edges" of the pulse) is known and equal to two. Thus, there is no need to consider the dimension changes for pulse locations. For simplicity, updates of the start and end points are random walk perturbation, performed by a local exploration of the posterior density $T^* \mid T \sim N(T, \sigma_T^2)$.

IV. RESULTS

A modified scanner (Sonos5500 Philips Medical Systems, Andover, MA, USA) is used to acquire echo signals from both linear scattering copper spheres and non-linear scattering MBs. The transmit pulses are 6-cycle sinusoidal signals with fundamental frequency from 1.28MHz to 3.7MHz. By de-convolving the received signals with the inverse response of receiver, the true frequency contents of echo signals can be achieved. As mentioned above, the filtered echo signal is modeled as a sum of sinusoids embedded in noise. Bayesian inference is then used to obtain the posterior density of the frequencies present, and numerical estimates are obtained using rjMCMC techniques. First, the performance of our algorithm is compared with a ground truth by analyzing synthetic data. Evaluated over 100 Monte Carlo runs, the percentage of correct estimation of number of frequencies is 91% and the average error of estimation of frequency values is just 0.01%. Then the algorithm is applied to real experimental signal generated under a transmit pulse with fundamental frequency f_0=1.28MHz and peak negative pressure at 300kPa.

A. *Analysis of linear solid sphere responses*

When the receiver characteristic is not taken into account, frequency estimation errors are present. For example, if f_0 lies in a region where the receiver attenuates the signal, f_0 cannot be correctly identified: the largest peak will be misidentified as the fundamental rather than a harmonic. Incorporating the receiver response into the analysis leads to correct detection of f_0.

B. *Analysis of non-linear MB responses:*

The FT spectrum of the MB signals is indistinct by observation even after additional optimization methods, such as thresholding, are combined to estimate the spectral peaks. The new proposed estimation algorithm can produce three frequencies at 2.52MHz, 3.13MHz and 3.78MHz. These are the 2nd, 2.5th and 3rd harmonics, consistent with the transmitted pulse. The convergence diagnostics is displayed in upper panel of Fig.7. From Fig. 7, after 1000 iterations, called the burn-in period, the estimate of the number of frequency converges. The lower panel of Fig.7 shows the probability of number of frequencies. The highest bar denotes the most probable number of frequencies. Furthermore, the pulse location is estimated accurately. Fig.6 depicts the original MB response in the time domain. The vertical line (659, 736) in the figure indicates the start and end points of the MB pulse segment. As the region of interest between the MB and receiver is between 6.8 and 8.2cms, this prior knowledge is used to provide an accurate estimate of the pulse location of 7.5cm. This helps extract the desired MB echo signal automatically for further analysis.

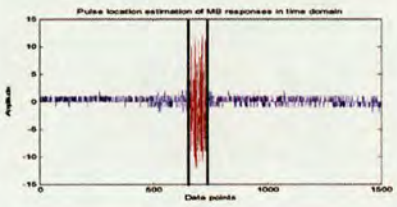


Fig.6. Original MB response in the time domain

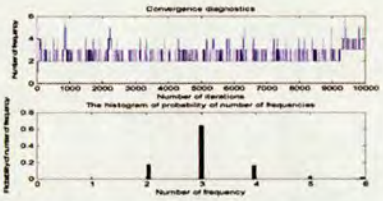


Fig.7. Convergence diagnostics and histogram of the number of frequency for single pulse MB response

V. CONCLUSIONS

This study presents a robust temporal and spectral estimation algorithm and accounts for the receiver response. The superiority of the new method over FT is demonstrated and the method will help with understanding MB behavior. This may also progress to the automatic classification of echo signals and is intended to be used for adaptive pulse sequence design.

REFERENCES

[1] E.Stride and N.Saffari, "Microbubble ultrasound contrast agents, a review," in proceedings of the Institute of Mechanical Engineers, Part H:Journal of Engineering in Medicine, vol.217, 2003.

[2] Gregory P. A Bayesian revolution in spectral analysis in Erickson GJ, Smith CR. "Bayesian inference and maximum entropy methods in schience and engineering. Kluwer," Academic publishers, USA, 2001

[3] H.Akaike, "A new look at the statistical model identification," in IEEE Trans. Automat. Contr., vol. AC-19, pp.716-723, 1974.

[4] J.Rissanen, "Modeling by shortest data description," in Automation, vol. 14, pp. 465–471, 1978.

[5] C.Andrieu and A.Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," in IEEE Trans. Signal Processing, vol 47, pp.2667-2676, 1999.

[6] V.Sboros, S.D.Pye, C.A.MacDonald, J.Gomatam, C.M.Moran and W.N.McDicken, "Absolute measurement of ultrasonic backscatter from single microbubles," Ultrasound in Medicine & Biology, vol.31, pp.1063-1072, 2005.

[7] Joseph J.K.O Ruanaidh and William J. Fitzgerald, "Numerical Bayesian methods applied to signal processing," Springer, 1996

[8] C.Andrieu, N.de Freitas, A. Doucet and M.I.Jordan, "An introduction to MCMC for machine learning," Machine learning, 50, 5-43, 2003

A NOVEL ESTIMATION SYSTEM FOR MULTIPLE PULSE ECHO SIGNALS FROM ULTRASOUND CONTRAST MICROBUBBLES

Yan Yan, James R.Hopgood and Vassilis Sboros *

Institute for Digital Communications,
School of Engineering and Electronics, University of Edinburgh, UK
{Y.Yan@ed.ac.uk, James.Hopgood@ed.ac.uk and vassilis.sboros@ed.ac.uk}

ABSTRACT

The understanding and exploitation of non-linear microbubble signals is an active research area that aims to advance contrast ultrasound into a high sensitivity and specificity diagnostic imaging modality. To discriminate the difference between echoes from tissue and contrast microbubbles, it is of significance to extract as much information of the reflected signals as possible, especially the pulse locations in the time domain and their corresponding spectral contents in the frequency domain. In this paper, a novel estimation system for extracting the information of interest is proposed. This estimation technique is based on non-parametric methods for coarse estimation, followed by a parametric method within Bayesian framework for estimation refinement. The results show that the pulse location and frequency content can be accurately estimated simultaneously. This assists in the design of transmit pulsing regimes in future work.

Index Terms— ultrasound contrast microbubbles, Bayesian inference, Monte Carlo methods, parametric model

1. INTRODUCTION

Microbubbles have been widely used as Ultrasound Contrast Agents (UCAs) in bio-medical research area since the 1990's [1]. They are composed of gas-filled encapsulated microspheres, usually with diameter below $7\mu m$, that can go through microcirculations in the human body. The microbubbles have a non-linear acoustic signature, as they are more compressible when exposed to an oscillating acoustic signal compared to soft tissue [2]. In order to design a transmit pulse that can maximise the difference between responses from microbubbles and tissue and increase the contrast-to-tissue ratio (CTR) [3], it is of particular interest to jointly extract the characteristics of responses in both the time and frequency domains.

Most traditional methods for frequency estimation in ultrasound are based on the Fourier transform (FT) methods [3]. However, due to the limitations of frequency resolution, the FT can not detect some frequencies that may have important physical meanings, or may provide false spectral contents. Moreover, it does not localize in time. Additionally, in the time domain, pulse locations and durations attract more attention. The Hilbert transform [4] is widely used for envelope detection. It may fail when the signal is embedded in noise. Moreover, it only operates in the time domain and cannot offer any frequency information about the signal. There is little information in the ultrasonic literature about the joint estimation of pulse locations and frequencies, which can be further developed, especially for multiple pulse echo signals.

In this paper, a new ultrasound contrast microbubble estimation system is proposed. The whole procedure is outlined in section 2. In section 3, coarse estimation for pulse locations and spectral content based on nonparametric methods is described. In section 4, optimisation of parameters using a parametric model within the Bayesian framework provides more accurate estimates. Performance evaluation for the synthetic signals and experimental signals are both shown in section 5. Section 6 concludes that the proposed estimation system can extract the information in both time and frequency domains simultaneously.

2. ULTRASOUND CONTRAST MICROBUBBLE ESTIMATION SYSTEM

The design of a complete estimation system is required to extract the information of echo signals from microbubbles automatically. The system can be divided into two parts, as displayed in Fig.1. The first part is coarse estimation for pulse locations and the frequency contents; the second part is to refine the estimation by using a parametric model within Bayesian framework. Details are described later in sections 3 and 4.

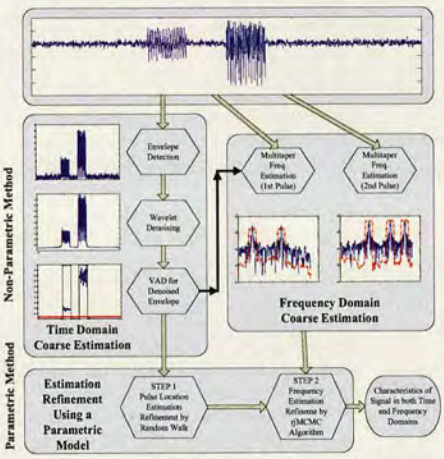


Fig. 1. Procedure for the whole estimation system

*Vassilis Sboros is with Department of Medical Physics. IDCOM is in the Joint Research Institute for Signal and Image Processing, a member of the Edinburgh Research Partnership in Engineering and Mathematics. This project is funded by Biologically Inspired Acoustic System (BIAS).

1-4244-1484-9/08/$25.00 ©2008 IEEE

3. COARSE ESTIMATION

3.1. Coarse Estimation in Time Domain

From the observation of responses from microbubbles, the techniques used for signal burst detection can be applied to pulse location estimation of the microbubble echo signals in ultrasound imaging. The estimation of pulse location is related to the estimation of the start and end points of each pulse.

3.1.1. Wavelet Denoising for Envelope Detection

Envelope detection using the Hilbert transform is widely used for changepoints (start and end points) estimation [4]. In the case of low signal-to-noise ratio(SNR), the envelope detected by the Hilbert transform is quite noisy, thus can not be easily extracted. In order to obtain a clearer envelope of the multiple pulse signal, wavelet denoising is adopted as it has more advantages compared with traditional filtering approaches, e.g. it is non-linear and can be applied to non-stationary signals. However, the method may fail when there are closely-spaced pulses present in the real bubble signal.

3.1.2. Voice Activity Detection

Voice activity detection (VAD) is an energy detector in various applications. The technique used by Alan [5] introduces a low-variance spectral estimator and determines an optimal threshold based on the estimated noise statistics. Nevertheless, if the SNR of the signal is low or the amplitudes of microbubble pulses are small, the performance of VAD is not acceptable since the threshold may not be properly chosen.

3.1.3. Coarse Estimation of Pulse Location

Although both the VAD algorithm and wavelet denoising for envelope detection have their own limitations, the combination of these two can improve the detection accuracy. Firstly, the Hilbert transform is used for envelope detection. Secondly, the envelope is denoised by a stationary wavelet transform, which is selected for its time-invariant property [6]. Finally, the VAD algorithm is performed for the denoised envelope. Following the aforementioned procedure of the algorithm, the proposed method can give better estimation for both closely-spaced pulses and small amplitude pulses. In addition, by experimental observation, the method can tolerate lower SNR down to $5dB$.

3.2. Coarse Estimation in Frequency Domain

The discrete Fourier transform (DFT) is often used to analyze the signal in the frequency domain. Andrieu and Doucet proposed a frequency estimation algorithm based on the DFT in [7]. This approach can sometimes overestimate the number of frequencies. Therefore, the multitaper spectral estimator is introduced to refine the sampling process in the algorithm. In this technique, several data windows are used on the same data record to obtain several modified periodograms. These periodograms are then averaged to produce the multitaper spectrum [8]. By reducing the variance, a much cleaner spectrum is achieved. Moreover, the bias and resolution loss can also be reduced for properly designed tapers.

4. ESTIMATION REFINEMENT

The aforementioned estimation techniques are all non-parametric methods, which can only provide coarse estimation of the parameters of interest. Furthermore, for frequency estimation, multitaper spectral estimator only works for one specific segment of pulse and cannot offer the number of frequencies automatically. When there are multiple pulses in the signal and the number of frequencies for each pulse is unknown, more advanced techniques, such as using a parametric model within Bayesian framework, can give more accurate estimates.

4.1. Signal Model

As the experimental transmit pulse in ultrasound appears to be composed of several cycles of a harmonic signal, the multiple pulse bubble echo can be modeled as several segments of sum of sinusoids. Assume there are m pulses in the observed signal with N data points. The multiple pulses model can be defined as follows: ($T_0 \triangleq 1$, $T_{2m+1} \triangleq N$ and $i = 0, \ldots, 2m$)

$$D_0 \quad : \quad x(t) = n(t)$$

$$D_{k_m} \quad : \quad x(t) = \begin{cases} n(t) & \text{if } T_i < t \leq T_{i+1} - 1, i \text{ is even} \\ x_i(t) + n(t) & \text{if } T_i < t \leq T_{i+1} - 1, i \text{ is odd} \end{cases}$$

where $x_i(t) = \sum_{j=1}^{k_i} a_{c_{j,k_i}} \cos(\omega_{j,k_i} t) + a_{s_{j,k_i}} \sin(\omega_{j,k_i} t)$.

The signal model can be written in a vector-matrix form:

$$\mathbf{x} = \mathbf{G}(\boldsymbol{\omega}_{k_m}, \boldsymbol{T}_{2m})\mathbf{a}_k + \mathbf{n} \tag{1}$$

where $\mathbf{n}$ is the zero-mean white Gaussian noise with variance σ_k^2 and

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \boldsymbol{G}_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \boldsymbol{G}_m \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The $\mathbf{G}$ matrix is of size $2N \times \sum_{i=1}^{m} k_i$. It contains the information about changepoints $(T_1, T_2, \ldots, T_{2m})$, which are related to the positions of multiple pulses in the echo signal. Each component $\boldsymbol{G}_j (j = 1, \ldots, m)$ in $\mathbf{G}$ matrix represents a single pulse, which has its own parameters. The number of sinusoids and other parameters $\boldsymbol{\theta}_k \triangleq (\boldsymbol{\omega}_k, \boldsymbol{a}_k, \sigma_k^2)$ are all unknown in each pulse. As far as each segment is concerned, the $\boldsymbol{G}_j$ matrix can be defined as:

$$\boldsymbol{G}_j = \begin{bmatrix} E(\omega_{k_1} t(T_{2j-1})) & \cdots & E(\omega_{k_m} t(T_{2j-1})) \\ E(\omega_{k_1} t(T_{2j-1} + 1)) & \cdots & E(\omega_{k_m} t(T_{2j-1} + 1)) \\ \vdots & \vdots & \vdots \\ E(\omega_{k_1} t(T_{2j} - 1)) & \cdots & E(\omega_{k_m} t(T_{2j} - 1)) \end{bmatrix}$$

where $E(\cdot) \triangleq [\cos(\cdot), \sin(\cdot)]$. Moreover, T_{2j-1} and T_{2j} are two corresponding changepoints at the start and end points respectively for each pulse segment. For each $\boldsymbol{G}_j$, k_m may have different values, which implies the different number of frequencies and their values in different pulses.

The likelihood function can be easily obtained according to the signal model (m in $\{k, \boldsymbol{\theta}_k\}_m$ represents different pulse segment):

$$p(\mathbf{x}|\{k, \boldsymbol{\theta}_k\}_m, \boldsymbol{T}_{2m}) = (2\pi\sigma_k^2)^{-N/2} \times \exp\left\{-\frac{1}{2\sigma_k^2} \parallel \mathbf{x} - \mathbf{G}(\boldsymbol{\omega}_{k_m}, \boldsymbol{T}_{2m})\boldsymbol{a}_k \parallel^2\right\}. \tag{2}$$

4.2. Reversible jump MCMC Algorithm for Frequency Estimation Refinement

According to Bayesian inference [9], the joint posterior distribution is achieved based on the properly chosen prior distribution of all the parameters. The joint prior distribution and the posterior distribution, after integrating out the nuisance parameters $\boldsymbol{a}_{k_m}$ and σ_k^2, are expressed as follows (more details can be found in [7] and v_0, γ_0, Λ, δ^2 are hyperparameters of the Bayesian model):

$$p(\{k,\boldsymbol{\theta}_k\}_m,\boldsymbol{T}_{2m}) = p(\boldsymbol{T}_{2m})p(\{k,\boldsymbol{a}_k,\omega_k\}_m|\sigma_k^2)p(\sigma_k^2)$$
$$\propto \frac{\Lambda^{k_m}}{k_m!}\exp(-\Lambda)\times\frac{1}{|2\pi\sigma_k^2\Sigma_{k_m}|^{1/2}}\times\frac{1}{\pi^{k_m}}\frac{1}{\sigma_k^2}$$
$$\times\exp[-\frac{\boldsymbol{a}_{k_m}^T\Sigma_{k_m}^{-1}\boldsymbol{a}_{k_m}}{2\sigma_k^2}]\times(\frac{1}{N-1}\frac{1}{N-2}\cdots\frac{1}{N-2m}) \quad (3)$$

where $\Sigma_{k_m}^{-1} = \delta^{-2}\mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})$.

$$p(\boldsymbol{T}_{2m},\{k,\boldsymbol{\omega}_k\}_m|\mathbf{x}) \propto (\gamma_0+\mathbf{x}^T\mathbf{P}_{k_m}\mathbf{x})^{-(N+v_0)/2}$$
$$\times\frac{(\Lambda/[(\delta^2+1)\pi])^{k_m}}{k_m!} \quad (4)$$

where $\mathbf{P}_{k_m} = \mathbf{I}_N - \mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{M}_{k_m}\mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})$, $\mathbf{I}_N$ is a $N\times N$ identity matrix and $\mathbf{M}_{k_m}^{-1} = \mathbf{G}^T(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m})\mathbf{G}(\boldsymbol{\omega}_{k_m},\boldsymbol{T}_{2m}) + \Sigma_{k_m}^{-1}$.

However, the joint posterior distribution is highly non-linear, which means the closed form of $p(\boldsymbol{T}_{2m},\{k,\boldsymbol{\omega}_k\}_m|\mathbf{x})$ can not be obtained. Therefore, a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm is introduced to sample from the complicated joint distribution and then to estimate the multiple pulse locations and frequency contents for each pulse simultaneously.

In the proposed algorithm, a multitaper spectrum is adopted as the proposal distribution to provide the initial guess for frequency estimation. The rjMCMC algorithm is then used to explore the regions around obvious peaks in the multitaper power spectrum. Furthermore, as the number of frequencies and their values are all unknown, the reversible jump MCMC technique is incorporated to select the model order of the frequency automatically.

4.3. Random Walk Update for Pulse Locations

Based on the initial guesses given by aforementioned combination algorithm of VAD and wavelet denoising for envelope detection, a random walk perturbation is adopted as the proposal distribution for refinement of the pulse location estimates. Specifically, the update of each changepoint depends on its previous value and performs a local exploration of the initial guess, which can be described as:

$$T^*|T \sim \mathcal{N}(T,\sigma_T^2). \quad (5)$$

where T and T^* are previous state and new state of the changepoint respectively. $\mathcal{N}(\cdot)$ represents the normal distribution with mean T and variance σ_T^2.

4.4. Refinement Algorithm Based on a Parametric Model

The refinement of the parameter estimates, for both pulse location and frequency content, using a parametric model with numerical Bayesian method, consists of two steps in each iteration. Firstly, the pulse locations are updated by random walk perturbation; secondly, after the estimation of pulse locations, frequency contents for each pulse can be updated using rjMCMC algorithm for a given specific set of changepoints. The procedure is summarized in Algorithm 1. Details of the birth, death and update moves can be found in [7].

Algorithm 1 rjMCMC Algorithm for Estimation Refinement

1: **Initialization:** set $(\{k,\boldsymbol{\theta}_k\}_m^{(0)}, \boldsymbol{T}_{2m}^{(0)})$.
2: **Iteration:**
3: **for** $i = 1$ to $numIteration$ **do**
4: Update each changepoint $\boldsymbol{T}_{2m}^{(i)}$ using random walk.
5: For each pulse segment m, update frequency contents:
6: a).Sample hyperparameters Λ and δ^2.
7: b).Sample u from $U_{(0,1)}$. (*uniform distribution*)
8: **if** $u \le b_{k_m^{(i)}}$ **then**
9: perform *birth move* of a new frequency
10: **else if** $(u \le b_{k_m^{(i)}} + d_{k_m^{(i)}})$ **then**
11: perform *death move* of an existing frequency
12: **else**
13: perform *update move* of a frequency randomly
14: **end if**
15: c).Sample nuisance parameters $\boldsymbol{a}_{k_m}$ and σ_k^2.
16: **end for**

Table 1. Comparison of Accuracy between Coarse Estimation and Estimation Refinement of Pulse Locations (start & end points) SNR = 5dB

True Locations	450	600	750	900
Coarse Estimation	460	600	780	890
$\lvert\triangle\varepsilon\rvert$	10	0	20	10
Estimation Refinement	448	598	750	900
$\lvert\triangle\varepsilon\rvert$	2	2	0	0

5. RESULTS AND EVALUATION

5.1. Evaluate the Estimation System on Synthetic Signals

According to the characteristics of the multiple pulse echo signal from ultrasound contrast microbubbles, the synthetic signal is simulated as several pulse segments with a sum of sinusoids in each segment. This synthetic signal is used to evaluate the performance of the whole estimation system.

The synthetic signal has 1500 data points, which consists of two pulses. The first pulse locates between (450, 600) and has two frequency components ($0.2\pi = 0.6283$, $0.6\pi = 1.8849$); the second one locates between (750, 900) and has three frequencies ($0.4\pi = 1.2566$, $0.8\pi = 2.5132$, $0.3\pi = 0.9425$). After the estimation procedure, the coarse estimation and the estimation refinement for pulse locations are compared in Table 1. The error is calculated as the sum of difference between true values and estimated values of each pulse location. The error after estimation refinement is much less than that of coarse estimation only. Moreover, the frequency contents can be estimated at the same time with high accuracy {(0.6288, 1.8846); (1.2560, 2.5130, 0.9433)} compared to multitaper estimation { (0.6250, 1.8860); (1.2552, 2.5145, 0.9440) }.

Furthermore, the above experiments has been repeated 100 times with different noise realizations and amplitudes and phase components. On average, the error of coarse estimation for pulse locations is 20 whereas the error after estimation refinement is 5. For frequency estimation, the error percentage of multitaper technique is around 1% and the error percentage of rjMCMC algorithm is about 0.01%. As a result, the new estimation system indicates its superi-

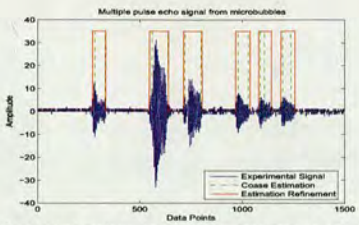


Fig. 2. Multiple pulses signal with pulse location estimation

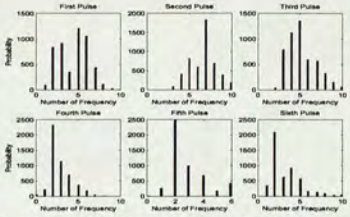


Fig. 4. Histogram of number of frequencies for each pulse segment

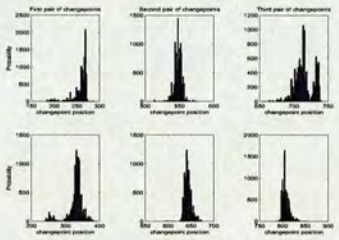


Fig. 3. Histogram of pulse locations for the first three pulses

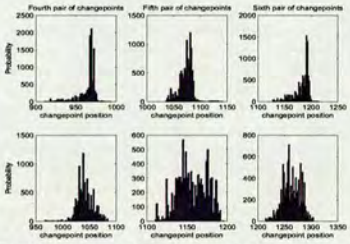


Fig. 5. Histogram of pulse locations for the last three pulses

ority in both the time domain and frequency domain.

5.2. Experimental Signal Analysis

In the following experiments, the microbubbles used for analysis in this paper are "Definity " [10], which are exposed to ultrasound peak negative pressure of 300kPa and transmit frequencies ranging from 1.1MHz to 3.2MHz. The raw signals produced from the microbubbles were preamplified, collected and stored. The multiple pulse signal shown in Fig.2 was collected with peak negative pressure of 300kPa and transmit frequency of 1.48MHz. Fig.3 and Fig.5 show histograms of the probability of pulse locations after estimation refinement {(269, 333), (548, 641), (715, 805), (969, 1040), (1082, 1145), (1191, 1259)}. The estimation result is also shown in Fig.2, denoting the locations for six pulses: The dash line represents the coarse estimation for pulse locations and the solid line represents the pulse location estimation after refinement, which offers more accurate estimates. For frequency estimation, Fig.4 depicts the histogram of number of frequency components in each pulse:(5, 7, 5, 2, 2, 2).

The whole algorithm has been carried out for other available data sets. The investigation indicates that the proposed estimation system can estimate the number of pulses in the signal and their positions, as well as the number of frequencies in each pulse and their corresponding values with very small errors simultaneously.

6. CONCLUSIONS

This paper proposes a novel estimation system for echo signals from ultrasound contrast microbubbles. The system first obtains coarse estimation by using non-parametric methods and then optimises the estimation by incorporating a parametric model within a Bayesian framework. The advantage is that it allows an automatic estimation of frequencies for each pulse and the pulse locations at the same time. Moreover, it exhibits improved frequency resolution compared to the Fourier analysis based techniques. Additionally, the parametric model introduced in the paper optimises the estimation of all parameters of interest. As a result, the new estimation system reveals more attributes in both time and frequency domains, which may broaden the research field in ultrasound contrast agents, especially in design of transmit pulsing regimes.

7. REFERENCES

[1] Gramiak R et al, "Echocardiography of the aortic root," *Invest Radio*, vol. 3, no. 356-366, 1968.

[2] E.Stride et al, "Microbubble ultrasound contrast agents, a review," *Eng. in Medicine*, vol. 217, 2003.

[3] Burns P.N. et al, "Higher harmonics of vibrating gas filled microspheres, part one, simulations," *Ultrasonics*, vol. 32, no. 447-453, 1994.

[4] Xander A et al, "Baseband velocity estimation for second-harmonic signals exploiting the invariance of the doppler equation," *IEEE Trans. on Bimedical Engineering*, vol. 45, no. 10, October 1998.

[5] A.Davis et al, "Statistical voice activity detection using low-variance spectrum estimation and an adaptive threshold," *IEEE Trans. on Speech and Audio Processing*, vol. 14, pp. 412–424, March 2006.

[6] X. Lu et al, "Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products," *Photogram. Eng. & Remote Sensing*, vol. 73, October 2007.

[7] C. Andrieu and A. Doucet, "Joint model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. on Signal Processing*, vol. 47, no. 10, October 1999.

[8] D. G.Manolakis, V. K.Ingle, and S. M.Kogon, *Statistical and adaptive signal processing:Spectral estimation, signal modelling, adaptive filtering and array processing*, McGrawHill, 1st edition, August 2003.

[9] P.C.Gregory, "A bayesian revolution in spectral analysis," in *Bayesian inference and maximum entropy methods in science and engineering*. Amer.Inst.of Physics Proceedings, 2001, pp. 557–568.

[10] V.Sboros et al, "The measurement of backscatter from individual contrast agent microbubbles," in *Ultrasonic Symposium*. IEEE, 2002, pp. 1945–1947.

Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

ThD01.4

Analysis of Echo Signal from Single Ultrasound Contrast Microbubble Using a reversible jump MCMC Algorithm

Yan Yan, James R. Hopgood and Vassilis Sboros

Abstract— **The understanding and exploitation of non-linear microbubble signals is an active research area that aims to advance contrast ultrasound into a high sensitivity and specificity diagnostic imaging modality. In order to discriminate the difference between echoes from tissue and contrast microbubbles, it is of particular interest to estimate the reflected signal pulse location in the time domain and its spectral content in the frequency domain. Therefore, a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm, a robust statistical signal processing technique, is introduced in this paper for the analysis of echo signals from Ultrasound Contrast Agents (UCAs). This algorithm provides many advantages over conventional Fourier transform based techniques. Furthermore, our results also show that the frequency components and pulse location can be accurately estimated simultaneously, which assists in characterising the signal content and the design of transmit pulsing regimes in future work.**

I. INTRODUCTION

The use of microbubbles as Ultrasound Contrast Agents (UCAs) was first discovered accidentally by cardiologist Charles Joiner in the 1960's [1]. It was not until the early 90's, with the production of stabilised biocompatible microbubbles, that UCAs started to be widely used in many research applications. UCAs are composed of gas-filled encapsulated microspheres small enough (usually with diameter between $2\mu m$ and $7\mu m$) to go through micro-circulations in the human body. Compared to soft tissue, the microbubbles are more compressible and expandable when exposed to an oscillating acoustic signal [2]. Therefore, the microbubbles have a non-linear acoustic signature.

The detection and estimation of echo signals from UCAs are of significant interest for diagnostic ultrasound, especially for differentiating the responses from microbubbles and those from tissue and increasing the Contrast-to-Tissue Ratio (CTR) [3]. Among many characteristics of echo signals from microbubbles and tissue, the frequency components and the pulse location in signal give most of the information for discriminating the echo signals. Accordingly, we focus on analyzing frequencies and pulse location simultaneously.

Methods based on the Fourier transform for frequency estimation are usually employed for the purpose that the ultrasound transmit pulses consist of several sinusoids. However, Akay pointed out some inherent limitations of techniques based on the Fourier transform in [4]. Firstly, the Fourier transform does not localize in time and, secondly, the limitations of frequency resolution may not detect some frequency components which can have important physical meanings. Moreover, envelope detection using the Hilbert transform is widely used for changepoint estimation [5]. Nevertheless, this method can not give accurate estimation if there are more than one dominant frequency in the signal. On the other hand, rjMCMC algorithm [6] is adopted to echo signal estimation to solve the aforementioned limitations.

In this paper, we introduce the rjMCMC algorithm to the analysis of echo signals from microbubbles. First, we describe the algorithm within the Bayesian framework and apply it to frequency estimation of the synthetic signal to show the advantages over Fourier analysis. Second, we apply rjMCMC to the real signal and it shows that the frequency components of the echo signal from microbubbles can be automatically detected and estimated. Finally, we combine the rjMCMC algorithm for frequency estimation with pulse location estimation (i.e. changepoint detection) in order to analyze the signal characteristics in time domain and frequency domain simultaneously.

II. FREQUENCY ESTIMATION USING RJMCMC ALGORITHM

In 1965, the Fast Fourier transform (FFT) was introduced [7] and then the discrete Fourier transform (DFT) became ubiquitous in spectral analysis. Based on the DFT, there are many other *non-parametric* techniques in spectral analysis, such as the periodogram which do not assume a particular functional form and have resolution limitations [8].

In 1987, Jaynes [7] applied the principles of Bayesian inference into the spectral analysis and found that the spectral resolution is better than conventional Fourier power spectrum or periodogram. Therefore, use of the FFT is not optimal and more information can be extracted from the data if more sophisticated statistics, such as prior knowledge of the signal and noise, are used. Additionally, after applying *non-parametric* techniques, methods to determine the accuracy of the peak frequencies are still needed whereas the Bayesian posterior probability distribution function for the signal frequency gives this information directly [7]. Therefore, Bayesian inference is introduced as the *parametric* technique, which proposes a signal model according to the characteristics of the real experimental signal.

A. Bayesian Inference

Considering that the experimental transmit pulse in ultrasound is composed of several cycles of sinusoids, we model

Yan Yan and James R. Hopgood are with the Institute of Digital Communications, School of Engineering and Electronics, University of Edinburgh, Edinburgh, U.K. { Y.Yan, James.Hopgood } @ed.ac.uk

Vassilis Sboros is with the Department of Medical Physics, School of Clinical Sciences and Community Health, University of Edinburgh, Edinburgh, U.K. Vassilis.Sboros@ed.ac.uk

1-4244-0788-5/07/$20.00 ©2007 IEEE

the one pulse echo signal from microbubbles as several superpositioned sinusoids and corrupted by Gaussian noise according to the non-linear property of microbubbles.

1) Data Model and Likelihood Function: Assume there are N observed data samples. T_1 & T_2 are two changepoints.

$$D_0 : x(i) = n(i)$$

$$D_k : x(i) = \begin{cases} n(i) & 0 \le i \le T_1 - 1 \\ \sum_{j=1}^{k}(a_{c_{j,k}} \cos(\omega_{j,k} i) + & \\ a_{s_{j,k}} \sin(\omega_{j,k} i)) + n(i) & T_1 \le i \le T_2 - 1 \\ n(i) & T_2 \le i \le N - 1 \end{cases}$$

where k denotes the number of sinusoids in the signal model, $a_{c_{j,k}}$, $a_{s_{j,k}}$ and $\omega_{j,k}$ are the amplitudes and the radial frequency of the j-th sinusoid for the model respectively. Moreover, T_1 and T_2 are two changepoints, which determine the location of signal pulse. It can also be written in vector-matrix form:

$$\mathbf{x} = \mathbf{G}(\boldsymbol{\omega}_k)\boldsymbol{a}_k + \mathbf{n} \tag{1}$$

$$\mathbf{G}(\boldsymbol{\omega}_k) = \begin{pmatrix} 0 \;\; 0 & \dots & 0 \;\; 0 \\ \vdots & \ddots & \vdots \\ 0 \;\; 0 & \dots & 0 \;\; 0 \\ E(\omega_1 t(T_1)) & \dots & E(\omega_k t(T_1)) \\ \vdots & \ddots & \vdots \\ E(\omega_1 t(T_2 - 1)) & \dots & E(\omega_k t(T_2 - 1)) \\ 0 \;\; 0 & \dots & 0 \;\; 0 \\ \vdots & \ddots & \vdots \\ 0 \;\; 0 & \dots & 0 \;\; 0 \end{pmatrix}$$

where the matrix $\mathbf{G}(\boldsymbol{\omega}_k)$ is of size $N \times 2k$ and $E(\cdot) \triangleq [\cos(\cdot), \sin(\cdot)]$, $\mathbf{n}$ is the zero-mean White Gaussian Noise(WGN) with variance σ_k^2. The number of sinusoids k and their parameters $\boldsymbol{\theta}_k \triangleq (\boldsymbol{\omega}_k, \boldsymbol{a}_k, \sigma_k^2)$ are all unknown. And our objective is to estimate k and $\boldsymbol{\theta}_k$, especially the frequency components $\boldsymbol{\omega}_k$. Therefore, from the signal model given above, the likelihood function, which is used to predict unknown outcomes based on known parameters, is defined:

$$p(\mathbf{x}|k, \boldsymbol{\theta}_k) = (2\pi\sigma_k^2)^{-N/2} \times \exp\{-\frac{1}{2\sigma_k^2}[\mathbf{x} - \mathbf{G}(\boldsymbol{\omega}_k)\boldsymbol{a}_k]^T \times [\mathbf{x} - \mathbf{G}(\boldsymbol{\omega}_k)\boldsymbol{a}_k]\} \tag{2}$$

2) Prior and Posterior Distributions: Following Bayesian inference, we may draw samples from the posterior distribution $p(k, \boldsymbol{\omega}_k|\mathbf{x})$ given the appropriate prior distributions. These priors reflect our degree of belief of the relevant values of the parameters.

The joint prior distribution for all the parameters and the posterior distribution after integrating out the nuisance parameters $\boldsymbol{a}_k$ and σ_k^2 are denoted as follows [9]: ($v_0, \gamma_0, \Lambda, \delta^2$ are hyperparameters of the Bayesian model)

$$p(k, \boldsymbol{\theta}_k) = p(k, \boldsymbol{a}_k, \boldsymbol{\omega}_k|\sigma_k^2)p(\sigma_k^2) \propto \frac{\Lambda^k}{k!} \times \exp(-\Lambda)\frac{1}{|2\pi\sigma_k^2\Sigma_k|^{1/2}} \times \exp[-\frac{\boldsymbol{a}_k^T\Sigma_k^{-1}\boldsymbol{a}_k}{2\sigma_k^2}]\frac{1}{\pi^k}\frac{1}{\sigma_k^2} \tag{3}$$

where $\Sigma_k^{-1} = \delta^{-2}\mathbf{G}^T(\boldsymbol{\omega}_k)\mathbf{G}(\boldsymbol{\omega}_k)$.

$$p(k, \boldsymbol{\omega}_k|\mathbf{x}) \propto (\gamma_0 + \mathbf{x}^T\mathbf{P}_k\mathbf{x})^{-(N+v_0)/2} \cdot \frac{(\Lambda/[(\delta^2+1)\pi])^k}{k!} \tag{4}$$

where $\mathbf{P}_k = \boldsymbol{I}_N - \mathbf{G}(\boldsymbol{\omega}_k)\mathbf{M}_k\mathbf{G}^T(\boldsymbol{\omega}_k)$, $\boldsymbol{I}_N$ is a $N \times N$ identity matrix and $\mathbf{M}_k^{-1} = \mathbf{G}^T(\boldsymbol{\omega}_k)\mathbf{G}(\boldsymbol{\omega}_k) + \Sigma_k^{-1}$.

Although the posterior distribution is simplified, it is still highly non-linear, which means $p(k|\mathbf{x})$ can not be obtained in closed form thus there is no easy way to find an exact solution. However, by using the rjMCMC algorithm these difficulties can be bypassed.

B. rjMCMC Algorithm

In standard MCMC algorithm, instead of sampling from the joint posterior distribution directly, we create an ergodic Markov chain ($k^{(i)}$, $\boldsymbol{\theta}_{k^{(i)}}$) whose equilibrium distribution is the specified joint posterior distribution given by (4) and run the simulation long enough to reach the stationary distribution.

Since the model order of the signal is unknown, the algorithm proposed, called rjMCMC, is able to jump between subspaces of different dimensions [6], which is equivalent to switching between different model orders. More specifically, Maximum A Posteriori (MAP) estimator is used to determine the value of k that maximizes $p(k|\boldsymbol{x})$. The probabilities for choosing birth, death and update moves, denoted as b_k, d_k and u_k respectively, are described in [9]. The main step of the rjMCMC algorithm is as follows:

Algorithm 1 Main Part of rjMCMC Algorithm

1: **Initialization:** set ($k^{(0)}$, $\boldsymbol{\theta}_k^{(0)}$).
2: **Iteration:**
3: **for** $i = 1$ to $numIteration$ **do**
4: Sample hyperparameters Λ and δ^2
5: Sample u from $U_{(0,1)}$ (*uniform distribution*)
6: **if** $u \le b_{k^{(i)}}$ **then**
7: perform *birth move*
8: **else if** ($u \le b_{k^{(i)}} + d_{k^{(i)}}$) **then**
9: perform *death move*
10: **else**
11: perform *update move*
12: **end if**
13: Sample nuisance parameters $\boldsymbol{a}_k$ and σ_k^2
14: **end for**

For *birth* and *death* move, the dimension of the parameters changes according to changes of the state of Markov chain. For further details, refer to [9].

For *update* move, iterations are performed for all k frequencies. With probability λ, (λ is a real number between 0 and 1), we perform an Metropolis-Hasting (MH) step with invariant distribution $p(\omega_{j,k} \mid \mathbf{x}, \boldsymbol{\omega}_{-j,k})$, ($\boldsymbol{\omega}_{-j,k} \triangleq [\omega_{1,k}, \dots, \omega_{j-1,k}, \omega_{j+1,k}, \dots]$) and proposal distribution $q_1(\boldsymbol{\omega}_{j,k}^* \mid \boldsymbol{\omega}_{j,k})$; Similarly, with probability $1 - \lambda$,

TABLE I
COMPARISON OF TRUE FREQUENCY AND ESTIMATED FREQUENCY

True Freq.	0.3967	0.4366	0.5225	0.5739	0.6625
Freq.(rjMCMC)	0.3968	0.4366	0.5226	0.5739	0.6625
$\triangle\varepsilon$ (rjMCMC)	0.25%	0	0.02%	0	0
Freq. (FFT)	0.3927	0.4388	0.5250	0.5740	0.6600
$\triangle\varepsilon$ (FFT)	1.00%	0.50%	0.48%	0.02%	0.38%
True Freq.	0.7002	0.7402	0.7771	0.8125	0.8466
Freq.(rjMCMC)	0.7001	0.7403	0.7771	0.8125	0.8467
$\triangle\varepsilon$ (rjMCMC)	0.01%	0.01%	0	0	0.01%
Freq. (FFT)	0.7035	0.7460	0.7770	0.8130	0.8475
$\triangle\varepsilon$ (FFT)	0.49%	0.78%	0.01%	0.06%	0.11%

we perform an MH step with invariant distribution $p(\omega_{j,k} \mid \mathbf{x}, \boldsymbol{\omega}_{-j,k})$ and proposal distribution $q_2(\omega^{\star}_{j,k} \mid \omega_{j,k})$.

In [9], Andrieu et. al. proposed the proposal distribution q_1 based on Fourier transform, which is independent of previous state and q_2 as random walk perturbation, which is a zero-mean Gaussian random variable with variance σ^2_{RW}, of the previous frequency: $\omega^{\star}_{j,k} \mid \omega_{j,k} \sim N(\omega_{j,k}, \sigma^2_{RW})$.

However, due to the false peaks that are usually observed in FFT as a result of leakage of sidelobes, it sometimes tends to overestimate the order of frequencies. Thus in our algorithm, to obtain more accurate frequency for initial guess, we use multitaper technique instead of the basic FFT method when proposing q_1. In multitaper technique [8], several data tapers (data windows) are used on the same data record to get several modified periodograms. Then these periodograms are averaged to produce the multitaper spectrum. Therefore, a much cleaner spectrum is obtained by averaging, which can reduce variance. Moreover, the bias and loss of resolution can be reduced for properly designed tapers. Thus the basic idea of choosing q_1 is to propose a new frequency, independent of the previous one, and explore the regions around obvious peaks in multitaper power spectrum. Furthermore, the acceptance probability for each frequency is given in [9].

III. CHANGEPOINT DETECTION FOR ECHO SIGNAL FROM MICROBUBBLE

In [10] Punskaya performed a changepoint detection using rjMCMC algorithm. However, in our single pulse echo signal, the number of changepoints is already known as two, so we do not need to consider the dimension changes for changepoints. Therefore, in order to simplify the problem, we update each new changepoint position by random walk perturbation, which performs a local exploration of the posterior distribution. Moreover, for the purpose of ensuring irreducibility of the Markov chain, the initial values of two changepoints are randomly chosen between 1 and $N-2$.

IV. EVALUATION AND RESULTS

A. Synthetic Signal Analysis using rjMCMC Algorithm

The echo signal from microbubbles is supposed to have several sinusoidal components, which may be close to each other, and be nonlinear to transmit pulse that has only one fundamental frequency. Therefore, according to characteristics of the echo signal, a synthetic signal simulated as a sum of sinusoids consisting of ten closely-spaced frequencies with varying amplitudes and phases is used to evaluate the rjMCMC algorithm.

In Table I, the error percentage of rjMCMC algorithm for each frequency estimation is lower than that of FFT , which indicates the superiority of rjMCMC algorithm. Moreover, the estimation of model order (number of frequencies) $k = 10$ is determined directly rather than introducing other optimization methods to find the peaks in spectrum like classic Fourier analysis techniques.

Furthermore, the above experiments has been repeated 100 times with different noise realizations and amplitude and phase components. We find out that on average, the error percentage of FFT based technique is around 1% and the error percentage of rjMCMC algorithm is about 0.01%. Accordingly, we can regard the rjMCMC algorithm as a robust tool for frequency estimation at the price of increased computational complexity compared to FFT.

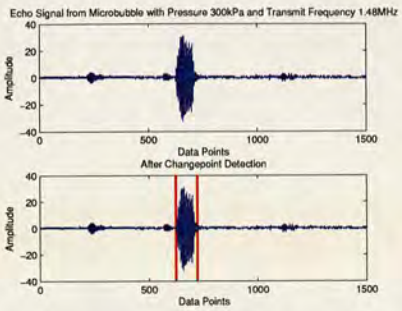


Fig. 1. Echoes from Bubble Signal in Time Domain

B. Frequency Estimation of Echo Signal from Single Microbubble

In our experiment, the echo signals from microbubbles used in this paper for analysis are Definity [11], which are exposed to ultrasound peak negative pressure of 300kPa. In order to analyze the characteristics of echo signal from microbubbles, we collected the echo signals under the transmit frequencies ranging from 1.1MHz to 3.2MHz and saved in text files as different data sets. The signal shown in Fig.1 was collected as scattering from microbubbles with peak negative pressure of 300kPa and transmit frequency of 1.48MHz.

Then we run 10000 iterations using rjMCMC for the frequency estimation of echo signal in Fig.1. From the lower part of Fig.2, a burn-in period [6] of first 4000 iterations is thrown away and the rest 6000 samples are drawn from the posterior distribution (shown by vertical line in the lower part of Fig.2). From the upper part of Fig.2, we can clearly see the model order (i.e.number of frequnecies) is $k = 9$.

These estimated frequencies offer more harmonics around

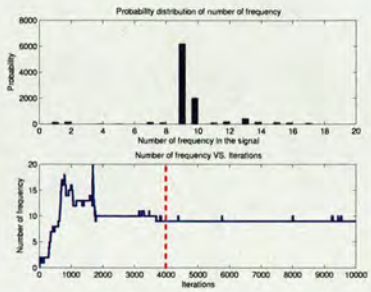


Fig. 2. Number of Frequency and the Convergence Diagnosis (1.5901, 1.9345, 2.3545, 2.5711, 2.7871, 2.9527, 3.1487, 3.5438, 3.7621 MHz)

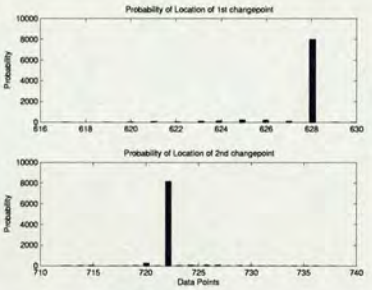


Fig. 3. Probability of Changepoint Locations using MAP Criterion

2^{nd} harmonic frequency. Therefore, more information of microbubble behavior can be extracted from the echo signal.

C. Changepoint Detection of Echo Signal from Single Microbubble

Fig.3 shows the probability of two changpoint locations and Fig.4 shows that the random walk perturbation for two changepoint locations converge quickly. Moreover, the two vertical lines in the lower part of Fig.1 denote the exact positions of changepoints $(628, 722)$ in experimental echo signal, which determines the pulse location of echo signal.

The same analysis was also carried out for other available data sets. The investigation indicates using rjMCMC algorithm can estimate the number of frequencies directly and the values of them with very small errors. At the same time, the pulse location can be estimated with frequency components simultaneously.

V. CONCLUSIONS AND FUTURE WORK

This paper introduces rjMCMC algorithm as a novel tool for analysis of echo signals from ultrasound contrast microbubbles in both time and frequency domain. The main advantage of this algorithm is that it allows an automatic estimation of frequencies in frequency domain and changepoints in time domain (relevant to pulse location)

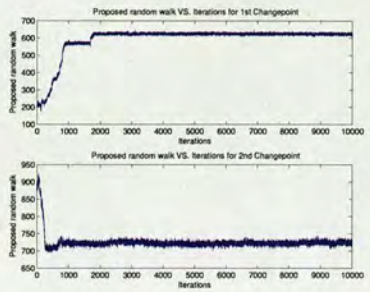


Fig. 4. Convergence Diagnosis: Changepoint Locations vs. Iterations

simultaneously. Moreover, it exhibits improved frequency resolution compared to the classical Fourier analysis. This algorithm presented here is a parametric algorithm that solves most limitations of Fourier transform, which is the standard technique to the study of echo signals from microbubbles. It is important to note that by using the rjMCMC algorithm, we are able to estimate, not only the 2^{nd} harmonics, but also more sub-harmonics and ultra-harmonics. As a result, it reveals more attributes in both time and frequency domain, which may broaden the research field in ultrasound contrast agents. Future work will be focused on estimation of multiple pulses echo signal from microbubbles and finally it will aim at the design of transmit pulsing regimes.

VI. ACKNOWLEDGEMENT

Authors give thanks to Prof. Steve McLaughlin, Dr. Yannis Kopsinis and Prof. Pradeep Misra for reviewing this manuscript. This work is supported by BIAS.

REFERENCES

[1] Gramiak R et al, Echocardiography of the aortic root, *Invest Radiol*, 3:356-366, 1968.
[2] E.Stride et al, Microbubble Ultrasound Contrast Agents: A Review, *Proc.Instn Mech.Engrs*, Vol.217, J.Engineering in Medicine, 2003.
[3] Burns P.N. et al, Higher Harmonics of Vibrating Gas Filled Microspheres: Part One:Simulations, *Ultrasonics*, 1994, 32: 447-453.
[4] M.Akay, Detection and Estimation Methods for Biomedical Signals, *San Diego-California-USA: Academic Press*, 1999.
[5] Xander A et al, Baseband Velocity Estimation for Second-Harmonic Signals Exploiting the Invariance of the Doppler Equation, *IEEE Trans. on Biomedical Engineering*, Vol.45, No.10, Oct.1998.
[6] J.Besage et al, Bayesian Computation and Stochastic Systems, *Stat.Sci.*, Vol.10, pp.3-66, 1995.
[7] P.C.Gregory, A Bayesian Revolution in Spectral Analysis, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris2000, Amer. Inst. of Physics Proceedings, 568, p.557, 2001.
[8] Dimitris G.Manolakis et al, Statistical And Adaptive Signal Processing: Spectral Estimation, Signal modelling, Adaptive Filtering and Array Processing, *The McGraw-Hill Companies,Inc*, 2000.
[9] Christophe Andrieu and Arnaud Doucet, Joint Model Selection and Estimation of Noisy Sinusoids via Reversible Jump MCMC, *IEEE Trans. on Signal Processing*, Vol.47, No.10, Oct,1999.
[10] Elena Punskaya et al, Bayesian Curve Fitting Using MCMC With Applications to Signal Segmentation, *IEEE Trans. on Signal Processing*, Vol.50,No.3,Mar.2002.
[11] V Sboros et al, The Measurement of Backscatter from Individual Contrast Agent Microbubbles, *IEEE Ultrasonics Symposium*, 1945-1947, 2002.